



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 148124

**TO: Janet Epps-Ford
Location: REM/2C05/2C18
Art Unit: 1635
Friday, March 18, 2005**

Case Serial Number: 08/901612

**From: David Schreiber
Location: Biotech-Chem Library
Remsen E01A61
Phone: 272-2526**

david.schreiber@uspto.gov

Search Notes

THIS PAGE LEFT BLANK

148124

Schreiber, David

From: Epps-Ford, Janet
Sent: Thursday, March 10, 2005 4:59 PM
To: Schreiber, David
Subject: sequence search request

Please search SEQ ID NOS: 59-65 of application 08/901,612, each sequence is under 30 nucleotides in length. Search all pending and published nucleic acid sequence databases.

Thanks,

Janet L. Epps-Ford, Ph.D.

Art Unit 1635

Mailbox: Remsen 2C18

Office: Remsen 2C05

Phone: 571-272-0757

Fax: 571-273-0757

THIS PAGE LEFT BLANK



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact **the searcher or contact:**

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



THIS PAGE LEFT BLANK

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY**Type of Search****Vendors and cost where applicable**

Searcher: <u>D. Schreder</u>	NA Sequence (#) <u>7</u>	STN _____
Searcher Phone #: <u>272-2526</u>	AA Sequence (#) _____	Dialog _____
Searcher Location: <u>Rensselaer EOL #61</u>	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: _____	Bibliographic _____	Dr. Link _____
Date Completed: <u>3/18</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: <u>15</u>	Fulltext _____	Sequence Systems <u>Comphgen</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time <u>10</u>	Other _____	Other (specify) _____

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 1025.6 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-59
Perfect score: 30
Sequence: 1 gacgaagaacagaagaauaggcagagt 30
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_uni:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	6	AR027810 Sequence
2	30	100.0	87	6	AX151115 Sequence
3	30	100.0	99	14	HPBPBPCA M76687 Hepatitis B
4	30	100.0	99	14	HPBPBPCA M76688 Hepatitis B
5	30	100.0	99	14	HPBPBPCA M76689 Hepatitis B
6	30	100.0	99	14	HPBPBPCA M76690 Hepatitis B
7	30	100.0	99	14	HPBPBPCA M76691 Hepatitis B
8	30	100.0	99	14	HPBPBPCA M76692 Hepatitis B
9	30	100.0	99	14	HPBPBPCA M76693 Hepatitis B
10	30	100.0	99	14	HPBPBPCA M76694 Hepatitis B
11	30	100.0	99	14	HPBPBPCA M76695 Hepatitis B
12	30	100.0	99	14	HPBPBPCA M76699 Hepatitis B
13	30	100.0	129	6	AX151114 Sequence
14	30	100.0	150	14	AF528205 Hepatitis B
15	30	100.0	150	14	AF528206 Hepatitis B
16	30	100.0	150	14	AF528207 Hepatitis B
17	30	100.0	150	14	AF528208 Hepatitis B
18	30	100.0	150	14	AF528209 Hepatitis B
19	30	100.0	150	14	AF528210 Hepatitis B

C 20	30	100.0	150	14	AF528211	Hepatitis
C 21	30	100.0	150	14	AF528212	Hepatitis
C 22	30	100.0	150	14	AF528213	Hepatitis
C 23	30	100.0	150	14	AF528214	Hepatitis
C 24	30	100.0	150	14	AF528215	Hepatitis
C 25	30	100.0	150	14	AF528216	Hepatitis
C 26	30	100.0	150	14	AF528217	Hepatitis
C 27	30	100.0	150	14	AF528218	Hepatitis
C 28	30	100.0	150	14	AF528219	Hepatitis
C 29	30	100.0	150	14	AF528220	Hepatitis
C 30	30	100.0	150	14	AF528221	Hepatitis
C 31	30	100.0	150	14	AF528222	Hepatitis
C 32	30	100.0	150	14	AF528224	Hepatitis
C 33	30	100.0	150	14	AF528225	Hepatitis
C 34	30	100.0	150	14	AF528226	Hepatitis
C 35	30	100.0	150	14	AF528227	Hepatitis
C 36	30	100.0	150	14	AF528228	Hepatitis
C 37	30	100.0	150	14	AF528229	Hepatitis
C 38	30	100.0	150	14	AF528231	Hepatitis
C 39	30	100.0	150	14	AF528232	Hepatitis
C 40	30	100.0	150	14	AF528233	Hepatitis
C 41	30	100.0	150	14	AF528234	Hepatitis
C 42	30	100.0	150	14	AF528235	Hepatitis
C 43	30	100.0	150	14	AF528236	Hepatitis
C 44	30	100.0	150	14	AF528237	Hepatitis
C 45	30	100.0	150	14	AF528238	Hepatitis

ALIGNMENTS

RESULT 1
AR027810
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN

Sequence 8 from patent US 5856459.
AR027810
AR027810.1 GI:5938630
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 30)
Frank, B.L., Roberts, P.C., Goodchild, J., Craig, J. Charles. and Mills, J.S.
Oligonucleotides specific for hepatitis B virus
Patent: US 5856459-A 8 05-JAN-1999;
Location/Qualifiers
1..30
/organism="unknown"
/mol_type="unassigned DNA"

Query Match
Best Local Similarity 100.0%; Score 30; DB 6; Length 30;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAGCAACAGAGAUUAGGCAGAGT 30
|||||
DB 1 GACATGCAACAGAGATGATTAGGCAGAGT 30

RESULT 2
AX151115/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

Sequence 4 from Patent WO0138498.
AX151115
AX151115.1 GI:14533317
synthetic construct
other sequences; artificial sequences.
1
Stuyver, L., Schinazi, R., de Gendt, S., van Geyt, C., Zoulim, F.,

```

Fried, M. and Roesau, R.
A new genotype of hepatitis B virus
Patent: WO 0138498-A 4 31-MAY-2001;
Pharmasset, Inc. (US); INNOGENETICS N.V. (BE)
FEATURES
    source
        1..87
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
ORIGIN
    Query Match      100.0%; Score 30; DB 6; Length 87;
    Best Local Similarity 86.7%; Pred. No. 0.0036;
    Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGCAUGAUUAGGCAGAGGT 30
    |||||:|||||:|||||:|||||:|||||
Db 43 GACATGAACAGAGATGATTAGGCAGAGGT 14
    |||||:|||||:|||||:|||||:|||||

RESULT 3
HPBPBREC/c
LOCUS
DEFINITION
    Hepatitis B virus type 1 precore protein (pre-C region, C) gene, 5'
end.
ACCESSION
    M76687
VERSION
    M76687.1 GI:485341
KEYWORDS
    e antigen; precore protein; tolerogen.
SOURCE
    Hepatitis B virus
ORGANISM
    Hepatitis B virus
REFERENCE
    1 Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
    Santantonio, T., Jung, M.C., Miska, S., Pastore, G., Pape, G.R. and
    Will, H.
    Prevalence and type of pre-C HBV mutants in anti-HBe positive
    carriers with chronic liver disease in a highly endemic area
    Virology 183 (2), 840-844 (1991)
91306476
MEDLINE
PUBMED
1853582
COMMENT
    Original source text: Hepatitis B virus DNA.
FEATURES
    source
        1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
        2
            /note="c in wt; t in virus type 2"
        10..93
            /gene="C"
        10..93
            /gene="C"
        /codon_start=1
        /product="precure protein"
        /protein_id="AAA45507.1"
        /db_xref="GI:485341"
        /translation="MQLFHLCLIIISCSPTVQASKLCGLWL"
        92
            /gene="C"
            /note="g in wt; a in virus type 2 (creates internal stop
            codon)"
ORIGIN
    Query Match      100.0%; Score 30; DB 14; Length 99;
    Best Local Similarity 86.7%; Pred. No. 0.0036;
    Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGCAUGAUUAGGCAGAGGT 30
    |||||:|||||:|||||:|||||:|||||
Db 52 GACATGAACAGAGATGATTAGGCAGAGGT 23
    |||||:|||||:|||||:|||||:|||||

RESULT 5
HPBPBREC/c
LOCUS
DEFINITION
    Hepatitis B virus type 3 precore protein (pre-C region, C) gene, 5'
end.
ACCESSION
    M76689
VERSION
    M76689.1 GI:485345
KEYWORDS
    e antigen; precore protein; tolerogen.
SOURCE
    Hepatitis B virus
ORGANISM
    Hepatitis B virus
REFERENCE
    1 Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
    Santantonio, T., Jung, M.C., Miska, S., Pastore, G., Pape, G.R. and
    Will, H.
    Prevalence and type of pre-C HBV mutants in anti-HBe positive
    carriers with chronic liver disease in a highly endemic area
    Virology 183 (2), 840-844 (1991)
91306476
MEDLINE
PUBMED
1853582
COMMENT
    Original source text: Hepatitis B virus DNA.
FEATURES
    source
        1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
        2
            /note="c in wt; t in virus type 2"
        10..93
            /gene="C"
        10..93
            /gene="C"
        /codon_start=1
        /product="precure protein"
        /protein_id="AAA45507.1"
        /db_xref="GI:485341"
        /translation="MQLFHLCLIIISCSPTVQASKLCGLWL"
        92
            /gene="C"
            /note="g in wt; a in virus type 1 (creates internal stop
            codon)"
ORIGIN
    Query Match      100.0%; Score 30; DB 14; Length 99;
    Best Local Similarity 86.7%; Pred. No. 0.0036;
    Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGCAUGAUUAGGCAGAGGT 30
    |||||:|||||:|||||:|||||:|||||
Db 52 GACATGAACAGAGATGATTAGGCAGAGGT 23
    |||||:|||||:|||||:|||||:|||||

RESULT 4
HPBPBREC/c
LOCUS
DEFINITION
    Hepatitis B virus type 2 precore protein (pre-C region, C) gene, 5'
end.
ACCESSION
    M76688
VERSION
    M76688.1 GI:485343
KEYWORDS
    e antigen; precore protein; tolerogen.
SOURCE
    Hepatitis B virus
ORGANISM
    Hepatitis B virus
REFERENCE
    1 Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
    Santantonio, T., Jung, M.C., Miska, S., Pastore, G., Pape, G.R. and
    Will, H.
    Prevalence and type of pre-C HBV mutants in anti-HBe positive
    carriers with chronic liver disease in a highly endemic area
    Virology 183 (2), 840-844 (1991)
91306476
MEDLINE
PUBMED
1853582
COMMENT
    Original source text: Hepatitis B virus DNA.
FEATURES
    source
        1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
        2
            /note="c in wt; t in virus type 2"
        10..93
            /gene="C"
        10..93
            /gene="C"
        /codon_start=1
        /product="precure protein"
        /protein_id="AAA45508.1"
        /db_xref="GI:485344"
        /translation="MQLFHLCLIIISCSPTVQASKLCGLWL"
        92
            /gene="C"
            /note="g in wt; a in virus type 2 (creates internal stop
            codon)"
ORIGIN
    Query Match      100.0%; Score 30; DB 14; Length 99;
    Best Local Similarity 86.7%; Pred. No. 0.0036;
    Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGCAUGAUUAGGCAGAGGT 30
    |||||:|||||:|||||:|||||:|||||
Db 52 GACATGAACAGAGATGATTAGGCAGAGGT 23
    |||||:|||||:|||||:|||||:|||||

```

```

variation
/db_xref="taxon:10407"
6
/note="c in wt; t in virus type 3"
10..93
/gene="C"
CDS
10..93
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precure protein"
/protein_id="AAA45509.1"
/db_xref="GI:485348"
/translation="MQLFHLCLIISCSCTVQASKLCGLWL"
58
/gene="C"
/note="g in wt; t in virus type 3 (val to phe)"
92
/gene="C"
/note="g in wt; a in virus type 3 (creates internal stop codon)"
ORIGIN
Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. NO. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUUAGGCAGAGGT 30
|||||:|||||:|||||:|||||:|||||
DB 52 GACATGAACAGAGATGATTAGGCAGAGGT 23

RESULT 6
HPBPRED/c
LOCUS
DEFINITION
Hepatitis B virus type 4 precure protein (pre-C region, C) gene, 5'
end.
ACCESSION M76690.1 GI:485347
VERSION
KEYWORDS e antigen; precure protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
source
1..99
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:10407"
10..93
/gene="C"
10..93
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precure protein"
/protein_id="AAA45510.1"
/db_xref="GI:485350"
/translation="MQLFHLCLIISCSCTVQPSKLCGLWL"
64..67
variation
/gene="C"
/note="gcc in wt; ccg in virus type 5 (ala to pro)"
92
variation
/gene="C"
/note="g in wt; a in virus type 5 (creates internal stop codon)"
95
variation
ORIGIN
Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. NO. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUUAGGCAGAGGT 30
|||||:|||||:|||||:|||||:|||||
DB 52 GACATGAACAGAGATGATTAGGCAGAGGT 23

RESULT 8
HPBPRED/c
LOCUS
DEFINITION
Hepatitis B virus type 6 precure protein (pre-C region, C) gene, 5'
end.
ACCESSION M76692.1 GI:485351
VERSION
KEYWORDS e antigen; precure protein; tolerogen.
SOURCE Hepatitis B virus

```

```

Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. NO. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUUAGGCAGAGGT 30
|||||:|||||:|||||:|||||:|||||
DB 52 GACATGAACAGAGATGATTAGGCAGAGGT 23

RESULT 7
HPBPRED/c
LOCUS
DEFINITION
Hepatitis B virus type 5 precure protein (pre-C region, C) gene, 5'
end.
ACCESSION M76691.1 GI:485349
VERSION
KEYWORDS e antigen; precure protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
source
1..99
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:10407"
10..93
/gene="C"
10..93
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precure protein"
/protein_id="AAA45511.1"
/db_xref="GI:485350"
/translation="MQLFHLCLIISCSCTVQPSKLCGLWL"
64..67
variation
/gene="C"
/note="gcc in wt; ccg in virus type 5 (ala to pro)"
92
variation
/gene="C"
/note="g in wt; a in virus type 5 (creates internal stop codon)"
95
variation
ORIGIN
Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. NO. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUUAGGCAGAGGT 30
|||||:|||||:|||||:|||||:|||||
DB 52 GACATGAACAGAGATGATTAGGCAGAGGT 23

RESULT 8
HPBPRED/c
LOCUS
DEFINITION
Hepatitis B virus type 6 precure protein (pre-C region, C) gene, 5'
end.
ACCESSION M76692.1 GI:485351
VERSION
KEYWORDS e antigen; precure protein; tolerogen.
SOURCE Hepatitis B virus

```

```

ORGANISM      Hepatitis B virus
REFERENCE      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
AUTHORS        1 (bases 1 to 99)
                Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
                Will,H.
TITLE          Prevalence and type of pre-C HBV mutants in anti-HBe positive
                carriers with chronic liver disease in a highly endemic area
JOURNAL        Virology 183 (2), 840-844 (1991)
MEDLINE        91306476
PUBMED         1853582
COMMENT        Original source text: Hepatitis B virus DNA.
FEATURES       Location/Qualifiers
                source
                1..99
                /organism="Hepatitis B virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10407"
                gene
                10..99
                /gene="C"
                misc_feature
                10..99
                /gene="C"
                /product="precore protein"
                /note="putative cds"
                variation
                11
                /gene="C"
                /note="t in wt; c in virus type 6 (loss of start codon)"

ORIGIN
Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAAGAGAUUAGGCAGAGGT 30
    |||:|||||:|||||:|||||:|||||
Db 52 GACATGAACAAGAGATGATTAGGCAGAGGT 23

RESULT 9
HPBPREGC/c      HPBPREGC
LOCUS            Hepatitis B virus type 7 precore protein (pre-C region, C) gene, 5'
DEFINITION      end.
ACCESSION        M76693
VERSION          M76693.1 GI:485352
KEYWORDS          e antigen; precore protein; tolerogen.
SOURCE            Hepatitis B virus
ORGANISM          Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE        1 (bases 1 to 99)
AUTHORS            Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
                Will,H.
TITLE            Prevalence and type of pre-C HBV mutants in anti-HBe positive
                carriers with chronic liver disease in a highly endemic area
JOURNAL          Virology 183 (2), 840-844 (1991)
MEDLINE          91306476
PUBMED           1853582
COMMENT          Original source text: Hepatitis B virus DNA.
FEATURES         Location/Qualifiers
                source
                1..99
                /organism="Hepatitis B virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10407"
                gene
                10..93
                /gene="C"
                misc_feature
                10..93
                /gene="C"
                /product="precore protein"
                /standard_name="pre-C region note: putative CDS"
                variation
                10
                /gene="C"
                /note="a in wt; t in virus type 7 (loss of start codon)"
                variation
                14
                /gene="C"
                /note="a in wt; g in virus type 7 (gln to arg)"

```

```

variation      92
                /gene="C"
                /note="g in wt; a in virus type 7 (creates internal stop
                codon)"

ORIGIN
Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAAGAGAUUAGGCAGAGGT 30
    |||:|||||:|||||:|||||:|||||
Db 52 GACATGAACAAGAGATGATTAGGCAGAGGT 23

RESULT 10
HPBPREGC/c      HPBPREGC
LOCUS            Hepatitis B virus type 8 precore protein (pre-C region, C) gene, 5'
DEFINITION      end.
ACCESSION        M76694
VERSION          M76694.1 GI:485353
KEYWORDS          e antigen; precore protein; tolerogen.
SOURCE            Hepatitis B virus
ORGANISM          Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE        1 (bases 1 to 99)
AUTHORS            Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
                Will,H.
TITLE            Prevalence and type of pre-C HBV mutants in anti-HBe positive
                carriers with chronic liver disease in a highly endemic area
JOURNAL          Virology 183 (2), 840-844 (1991)
MEDLINE          91306476
PUBMED           1853582
COMMENT          Original source text: Hepatitis B virus DNA.
FEATURES         Location/Qualifiers
                source
                1..99
                /organism="Hepatitis B virus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10407"
                gene
                10..93
                /gene="C"
                misc_feature
                10..93
                /gene="C"
                /product="precore protein"
                /standard_name="pre-C region note: putative CDS"
                variation
                12
                /gene="C"
                /note="g in wt; t in virus type 8 (loss of start codon)"
                variation
                92
                /gene="C"
                /note="g in wt; a in virus type 8 (creates internal stop
                codon)"
                variation
                95
                /note="g in wt; a in virus type 8 (gly to asp)"

ORIGIN
Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAAAGAGAUUAGGCAGAGGT 30
    |||:|||||:|||||:|||||:|||||
Db 52 GACATGAACAAGAGATGATTAGGCAGAGGT 23

RESULT 11
HPBPREGC/c      HPBPREGC
LOCUS            Hepatitis B virus type 9 precore protein (pre-C region, C) gene, 5'
DEFINITION      end.
ACCESSION        M76695
VERSION          M76695.1 GI:485354

```

```
KEYWORDS      e antigen; precore protein; tolerogen.
SOURCE        Hepatitis B virus
ORGANISM      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE     1 (bases 1 to 99)
AUTHORS       Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
              Will,H.
TITLE         Prevalence and type of pre-C HBV mutants in anti-HBe positive
              carriers with chronic liver disease in a highly endemic area
JOURNAL       Virology 183 (2), 840-844 (1991)
MEDLINE       91306476
PUBMED        1853582
COMMENT       Original source text: Hepatitis B virus DNA.
FEATURES      Location/Qualifiers
source        1..99
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10407"
gene          10..99
              /gene="C"
misc_feature  10..99
              /gene="C"
              /product="precure protein"
              /standard_name="pre-C region note: putative CDS"
variation     13
              /gene="C"
              /notes="c in wt; t in virus type 9 (creates internal stop
              codon)"
variation     92
              /gene="C"
              /notes="g in wt; a in virus type 9 (creates internal stop
              codon)"
variation     95
              /notes="g in wt; a in virus type 9 (gly to asp)"

ORIGIN
Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGUAGGACAGAGGT 30
    |||||:|||||:|||||:|||||:|||||
Db 52 GACATGACACAGAGATGATTAGGACAGGT 23

RESULT 13
AX151114/c
LOCUS            AX151114                129 bp      DNA      linear      PAT 22-JUN-2001
DEFINITION       Sequence 3 from Patent WO0138498.
ACCESSION        AX151114
VERSION          AX151114.1
KEYWORDS         synthetic construct
SOURCE           synthetic construct
ORGANISM         other sequences; artificial sequences.
REFERENCE        1
AUTHORS          Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,
              Fried,M. and Rossau,R.
TITLE            A new genotype of hepatitis b virus
JOURNAL          Patent: WO 0138498-A 3 31-MAY-2001;
              Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES         Location/Qualifiers
source           1..129
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"

ORIGIN
Query Match      100.0%; Score 30; DB 6; Length 129;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGUAGGACAGAGGT 30
    |||||:|||||:|||||:|||||:|||||
Db 43 GACATGACACAGAGATGATTAGGACAGGT 14

RESULT 14
AF528205/c
LOCUS            AF528205                150 bp      DNA      linear      VRL 31-JUL-2003
DEFINITION       Hepatitis B virus ASC1123 core antigen precursor, gene, partial
              cds.
ACCESSION        AF528205
VERSION          AF528205.1
KEYWORDS         Hepatitis B virus
SOURCE           Hepatitis B virus
ORGANISM         Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE        1 (bases 1 to 150)
AUTHORS          Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
TITLE            Comparative evaluation of HBV precore and basal core promoter
              mutants in Indian patients with diverse clinical manifestations
              Unpublished
JOURNAL          2 (bases 1 to 150)
AUTHORS          Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
TITLE            Direct Submission
JOURNAL          Submitted (11-JUL-2002) Hepatitis Division, National Institute of
```

```
KEYWORDS      e antigen; precore protein; tolerogen.
SOURCE        Hepatitis B virus
ORGANISM      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE     1 (bases 1 to 99)
AUTHORS       Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
              Will,H.
TITLE         Prevalence and type of pre-C HBV mutants in anti-HBe positive
              carriers with chronic liver disease in a highly endemic area
JOURNAL       Virology 183 (2), 840-844 (1991)
MEDLINE       91306476
PUBMED        1853582
COMMENT       Original source text: Hepatitis B virus DNA.
FEATURES      Location/Qualifiers
source        1..99
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10407"
gene          10..99
              /gene="C"
misc_feature  10..99
              /gene="C"
              /product="precure protein"
              /standard_name="pre-C region note: putative CDS"
variation     13
              /gene="C"
              /notes="c in wt; t in virus type 9 (creates internal stop
              codon)"
variation     92
              /gene="C"
              /notes="g in wt; a in virus type 9 (creates internal stop
              codon)"
variation     95
              /notes="g in wt; a in virus type 9 (gly to asp)"

ORIGIN
Query Match      100.0%; Score 30; DB 14; Length 99;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGUAGGACAGAGGT 30
    |||||:|||||:|||||:|||||:|||||
Db 52 GACATGACACAGAGATGATTAGGACAGGT 23

RESULT 12
HPBPRECM/c
LOCUS            HPBPRECM                99 bp      DNA      linear      VRL 11-MAY-1994
DEFINITION       Hepatitis B virus type 13 precore protein (pre-C region, C) gene,
              5' end.
ACCESSION        M76699
VERSION          M76699.1
KEYWORDS         e antigen; precore protein; tolerogen.
SOURCE           Hepatitis B virus
ORGANISM         Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE        1 (bases 1 to 99)
AUTHORS          Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
              Will,H.
TITLE            Prevalence and type of pre-C HBV mutants in anti-HBe positive
              carriers with chronic liver disease in a highly endemic area
JOURNAL       Virology 183 (2), 840-844 (1991)
MEDLINE       91306476
PUBMED        1853582
COMMENT       Original source text: Hepatitis B virus DNA.
FEATURES      Location/Qualifiers
source        1..99
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10407"
gene          10..99
              /gene="C"
CDS           10..>99
```

Virology, 20-A, Dr Ambedkar Road, Pune, Maharashtra 411001, India

FEATURES

source
1..150
/organism="Hepatitis B virus"
/proviral
/mol_type="genomic DNA"
/isolate="ASC1123"
/isolation_source="asymptomatic HBsAg carrier"
/specific_host="Homo sapiens"
/db_xref="taxon:10407"
/country="India"
misc_feature
1..150
/note="contains partial basal core promoter"
64..150
/note="contains complete precore region"
/codon_start=1
/product="core antigen precursor"
/protein_id="AAP87556.1"
/db_xref="GI:32810972"
/translation="MQLFHLCLIIISCSCTVQASKLCLGLXG"
ORIGIN

Query Match 100.0%; Score 30; DB 14; Length 150;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACAUGAACAGAGAGAUUAGGCAGAGGT 30
|||||:|||||:|||||:|||||:|||||:|||||
Db 106 GACATGAACAAGAGATGATTAGGCAGAGGT 77

RESULT 15
AF528206/c
LOCUS
DEFINITION Hepatitis B virus ASC1112 core antigen precursor, gene, partial cds.
VERSION AF528206
KEYWORDS AF528206.1 GI:32810973
SOURCE
ORGANISM Hepatitis B virus
REFERENCE
1 Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
AUTHORS 1 (bases 1 to 150)
TITLE Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
Comparative evaluation of HBV precore and basal core promoter mutants in Indian patients with diverse clinical manifestations
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 150)
AUTHORS Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-2002) Hepatitis Division, National Institute of Virology, 20-A, Dr Ambedkar Road, Pune, Maharashtra 411001, India

FEATURES

source
1..150
/organism="Hepatitis B virus"
/proviral
/mol_type="genomic DNA"
/isolate="ASC1112"
/isolation_source="asymptomatic HBsAg carrier"
/specific_host="Homo sapiens"
/db_xref="taxon:10407"
/country="India"
misc_feature
1..150
/note="contains partial basal core promoter"
64..150
/note="contains complete precore region"
/codon_start=1
/product="core antigen precursor"
/protein_id="AAP87557.1"
/db_xref="GI:32810974"
/translation="MQLFHLCLIIISCSCTVQASKLCLGLXG"
ORIGIN

Query Match 100.0%; Score 30; DB 14; Length 150;
Best Local Similarity 86.7%; Pred. No. 0.0036;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAGAGAGAUUAGGCAGAGGT 30
|||||:|||||:|||||:|||||:|||||:|||||
Db 106 GACATGAACAAGAGATGATTAGGCAGAGGT 77

Search completed: March 17, 2005, 08:14:15
Job time : 1025.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 257 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-59
Perfect score: 30
Sequence: 1 gacgaacaagagagaauaggcagaggt 30

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn19808:*

2: Geneseqn19908:*

3: Geneseqn20008:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	2 AAT72562	Aat72562 Hepatitis
2	30	100.0	30	2 AAT72563	Aat72563 Hepatitis
3	30	100.0	39	10 ADC64742	Adc64742 Hepatitis
4	30	100.0	87	4 AAD09094	Aad09094 Hepatitis
5	30	100.0	129	4 AAD09093	Aad09093 Hepatitis
6	30	100.0	639	6 AAD27422	Aad27422 Hepatitis
7	30	100.0	639	6 AAD31509	Aad31509 Hepatitis
8	30	100.0	655	4 AAH77569	Aah77569 HBV genot
9	30	100.0	655	4 AAH77568	Aah77568 HBV genot
10	30	100.0	655	4 AAH77574	Aah77574 HBV genot
11	30	100.0	655	4 AAH77573	Aah77573 HBV genot
12	30	100.0	655	4 AAH77570	Aah77570 HBV genot
13	30	100.0	655	4 AAH77571	Aah77571 HBV genot
14	30	100.0	664	4 AAH77572	Aah77572 HBV genot
15	30	100.0	669	12 AD007220	Ado07220 Hepatitis
16	30	100.0	673	4 AAD09092	Aad09092 Hepatitis
17	30	100.0	673	4 AAH77563	Aah77563 HBV preCo
18	30	100.0	681	4 AAH77567	Aah77567 HBV genot
19	30	100.0	1395	2 AAV82688	Aav82688 Fulminant
20	30	100.0	1400	2 AAV82687	Aav82687 Fulminant

C 21	30	100.0	1445	2 AAV82692	Aav82692 Fulminant
C 22	30	100.0	1445	2 AAV82685	Aav82685 Fulminant
C 23	30	100.0	1445	2 AAV82690	Aav82690 Fulminant
C 24	30	100.0	1445	2 AAV82684	Aav82684 Fulminant
C 25	30	100.0	1500	2 AAV82695	Aav82695 Fulminant
C 26	30	100.0	1500	2 AAV82683	Aav82683 Fulminant
C 27	30	100.0	1500	2 AAV82694	Aav82694 Fulminant
C 28	30	100.0	1500	2 AAV82686	Aav82686 Fulminant
C 29	30	100.0	1500	2 AAV82706	Aav82706 Wild type
C 30	30	100.0	1500	2 AAV82689	Aav82689 Fulminant
C 31	30	100.0	1500	2 AAV82693	Aav82693 Fulminant
C 32	30	100.0	2342	1 AAN93072	Aan93072 Sequence
C 33	30	100.0	2743	1 AAN00003	Aan00003 Sequence
C 34	30	100.0	2743	2 AAQ04799	AAq04799 Recombina
C 35	30	100.0	3180	4 AAH42375	Aah42375 Nucleotid
C 36	30	100.0	3182	6 AAD31765	Aad31765 Hepatitis
C 37	30	100.0	3182	9 ACA62422	AcA62422 Hepatitis
C 38	30	100.0	3182	10 AAD60866	Aad60866 Hepatitis
C 39	30	100.0	3220	3 AAZ88924	Aaz88924 Hepatitis
C 40	30	100.0	3248	4 AAD09091	Aad09091 Hepatitis
C 41	30	100.0	3248	4 AAH77562	Aah77562 HBV genot
C 42	30	100.0	5618	2 AAQ88310	AAq88310 Plasmid p
C 43	30	100.0	7991	6 AAS16094	Aas16094 HBV viral
C 44	30	100.0	8007	6 AAS16092	Aas16092 HBV viral
C 45	30	100.0	8717	6 AAS16093	Aas16093 HBV viral

ALIGNMENTS

RESULT 1
AAT72562
ID AAT72562 standard; DNA; 30 BP.
XX
AC AAT72562;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV88b.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..30
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
XX
XX W09639502-A1.
XX
XX 12-DEC-1996.
XX
XX 04-JUN-1996; 96WO-EP002432.
XX
XX 06-JUN-1995; 95US-00467397.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
XX (HYBR-) HYBRIDON INC.
XX
XX Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
XX
XX Roberts NA, Roberts PC, Slade A;
XX
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
XX
XX used in the detection and treatment of HBV infection.
XX
XX Claim 1; Page 12; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV88b which
XX
XX is complementary to a portion of the hepatitis B virus (HBV) RNA. The
XX
XX antisense oligonucleotide may be used to detect the presence of HBV in a

CC sample. The antisense oligonucleotide, and oligonucleotides containing a
 CC sequence which is complementary to at least two non- contiguous regions
 CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
 CC cell or for the treatment of HBV infection

XX SQ Sequence 30 BP; 12 A; 3 C; 10 G; 5 T; 0 U; 0 Other;
 Query Match 100.0%; Score 30; DB 2; Length 30;
 Best Local Similarity 86.7%; Pred. No. 0.003; 0; Indels 0; Gaps 0;
 Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUACAAGAGAGAUUAGGCAGAGGT 30
 |||||:|||||:|||||:|||||:|||||
 Db 1 GACATGAACAAGAGATGATTAGGCAGAGGT 30

RESULT 2

AAT72563
 ID AAT72563 standard; DNA; 30 BP.

XX AC AAT72563;

XX 03-SEP-1997 (first entry)

XX Hepatitis B virus RNA antisense oligonucleotide HBV88Mb.

XX HBV; HBV infection; inhibition; replication; ss.

XX Synthetic.

Key	Location/Qualifiers
FT misc_feature	1..30
FT	/*tag= a
FT	/note= "Internucleotide linkages are phosphorothioate"
FT misc_RNA	1..20
FT	/*tag= b
FT	/note= "2'-OMe RNA"
FT modified_base	1
FT	/*tag= c
FT	/mod_base= gm
FT modified_base	2
FT	/*tag= d
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	3
FT	/*tag= e
FT	/mod_base= cm
FT modified_base	4
FT	/*tag= f
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	5
FT	/*tag= g
FT	/mod_base= um
FT modified_base	6
FT	/*tag= h
FT	/mod_base= gm
FT modified_base	7
FT	/*tag= i
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	8
FT	/*tag= j
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	9
FT	/*tag= k
FT	/mod_base= cm
FT modified_base	10
FT	/*tag= l
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	11

FT	/*tag= m
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	12
FT	/*tag= n
FT	/mod_base= gm
FT modified_base	13
FT	/*tag= o
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	14
FT	/*tag= p
FT	/mod_base= gm
FT modified_base	15
FT	/*tag= q
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	16
FT	/*tag= r
FT	/mod_base= um
FT modified_base	17
FT	/*tag= s
FT	/mod_base= gm
FT modified_base	18
FT	/*tag= s
FT	/mod_base= OTHER
FT	/note= "2'-O-methyladenosine"
FT modified_base	19
FT	/*tag= u
FT	/mod_base= um
FT modified_base	20
FT	/*tag= v
FT	/mod_base= um

WO9639502-A1.

12-DEC-1996.

04-JUN-1996; 96WO-EP002432.

06-JUN-1995; 95US-00467397.

(HOFF) HOFFMANN LA ROCHE & CO AG F.
 (HYBR-) HYBRIDON INC.

Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
 Roberts NA, Roberts PC, Slade A;

WPI; 1997-043124/04.

Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 used in the detection and treatment of HBV infection.

Claim 1; Page 12; 81pp; English.

The present sequence represents a synthetic oligonucleotide HBV88Mb which
 is complementary to a portion of the hepatitis B virus (HBV) RNA. The
 antisense oligonucleotide may be used to detect the presence of HBV in a
 sample. The antisense oligonucleotide, and oligonucleotides containing a
 sequence which is complementary to at least two non- contiguous regions
 of an HBV nucleic acid, may be used for inhibiting HBV replication in a
 cell or for the treatment of HBV infection

Sequence 30 BP; 12 A; 3 C; 10 G; 5 T; 4 U; 0 Other;

Query Match 100.0%; Score 30; DB 2; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.003;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAGAGAGAUUAGGCAGAGGT 30

Db 1 GACAUGAACAGAGAGAUUAGGCAGAGGT 30

RESULT 3
 ADC64742/c
 ID ADC64742 standard; RNA; 39 BP.
 XX
 AC ADC64742;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Hepatitis B virus DNA polymerase related RNA oligonucleotide.
 XX
 KW screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
 XX
 OS Synthetic.
 OS Hepatitis B virus.
 XX
 PN KR2002007891-A.
 XX
 PD 29-JAN-2002.
 XX
 PF 19-JUL-2000; 2000KR-00041420.
 XX
 PR 19-JUL-2000; 2000KR-00041420.
 XX
 PA (MOGA-) MOGAM BIOTECHNOLOGY INST.
 PA (VIRO-) VIROGEN CO LTD.
 XX
 PI Ji HJ, Jung SI, Kim YC, Min MG, Ryu WS, Yoon GS;
 XX
 DR WPI; 2003-309015/30.
 XX
 XX Screening of antiviral agents by protein-priming activity of hepatitis B virus DNA polymerase.
 XX
 PS Disclosure; Page 12; 13pp; Korean.
 XX
 CC The present invention describes a method of screening for an antiviral agent by the protein-priming activity of hepatitis B virus (HBV) DNA polymerase. Also described is developing an antiviral agent with a high selectivity to HBV which can be used for high-throughput screening. The present sequence represents an RNA oligonucleotide which is used in the CC exemplification of the present invention.
 XX
 SQ Sequence 39 BP; 5 A; 13 C; 3 G; 0 T; 18 U; 0 Other;
 Query Match 100.0%; Score 30; DB 10; Length 39;
 Best Local Similarity 86.7%; Pred. No. 0.0031; Mismatches 0; Indels 0; Gaps 0;
 Matches 26; Conservative 4;
 QY 1 GACAUGAACAGAGAGAUUAGGCAGAGGT 30
 Db 37 GACATGACACAGAGATGATTAGGCAGAGGT 8
 RESULT 4
 AAD09094/c
 ID AAD09094 standard; DNA; 87 BP.
 XX
 AC AAD09094;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Hepatitis B virus FR1 strain genotype G HBeAg DNA fragment.
 XX
 KW HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
 KW preS2; surface antigen; HBsAg; HBx protein; vaccine; HBeAg;
 KW liver disease; hepatitis; liver cancer; HBcAg; core antigen; ds.
 XX
 OS Hepatitis B virus.
 XX
 PN WO200138498-A2.
 XX
 PD 31-MAY-2001.

XX
 PF 21-NOV-2000; 2000WO-US032108.
 XX
 PR 24-NOV-1999; 99US-0167206P.
 XX
 PA (PHAR-) PHARMASSET INC.
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M;
 PI Rosseau R;
 XX
 DR WPI; 2001-367676/38.
 XX
 PT Novel hepatitis B virus genotype G, nucleic acids encoding virus,
 PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
 PT treat or prevent the hepatitis B virus genotype G infection in a subject.
 XX
 PS Claim 6; Page 57; 84pp; English.
 XX
 CC The present invention relates to hepatitis B virus (HBV) strain FR1,
 CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
 CC PreS2 and surface antigen HBsAg) and HBx proteins. HBV genotype G nucleic
 CC acids and polypeptides are useful for diagnosing, prognosing and treating
 CC infections caused by HBV genotype G. They can be used in a vaccine to
 CC treat or prevent HBV genotype G infection. The HBV genotype G derived
 CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
 CC sample or diagnosis of HBV genotype G infection. The presence of HBV
 CC genotype G statistically correlates with the presence of liver damage
 CC and/or liver cancer in the subject. The HBV genotype G core insert
 CC peptide encoding nucleic acid is useful for designing monitoring assays
 CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
 CC and HBeAg (genotype G e antigen) in patients infected with HBV. The
 CC antibodies or antigens of HBV genotype G are useful for identifying a
 CC stage of liver disease caused by HBV genotype G. The present sequence is
 CC a hepatitis B virus (HBV) strain FR1, genotype G DNA fragment encoding e
 CC antigen (HBeAg)
 XX
 SQ Sequence 87 BP; 14 A; 24 C; 17 G; 32 T; 0 U; 0 Other;
 Query Match 100.0%; Score 30; DB 4; Length 87;
 Best Local Similarity 86.7%; Pred. No. 0.0035; Mismatches 0; Indels 0; Gaps 0;
 Matches 26; Conservative 4;
 QY 1 GACAUGAACAGAGAGAUUAGGCAGAGGT 30
 Db 43 GACATGACACAGAGATGATTAGGCAGAGGT 14
 RESULT 5
 AAD09093/c
 ID AAD09093 standard; DNA; 129 BP.
 XX
 AC AAD09093;
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Hepatitis B virus FR1 strain genotype G DNA fragment #1.
 XX
 KW HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
 KW preS2; surface antigen; HBsAg; HBx protein; vaccine; liver disease;
 KW hepatitis; liver cancer; HBcAg; core antigen; ds.
 XX
 OS Hepatitis B virus.
 XX
 PN WO200138498-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 21-NOV-2000; 2000WO-US032108.
 XX
 PR 24-NOV-1999; 99US-0167206P.
 XX
 PD (PHAR-) PHARMASSET INC.

```

PA (INNO-) INNOGENETICS NV.
XX
XX Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M;
PI Rosau R;
XX
XX WPI; 2001-367676/38.
DR
XX
XX Novel hepatitis B virus genotype G, nucleic acids encoding virus,
PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX
XX Claim 5; Page 57; 84pp; English.
PS
XX
XX The present invention relates to hepatitis B virus (HBV) strain FRI,
CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
CC PreS2 and surface antigen HBsAg) and HBx proteins. HBV genotype G nucleic
CC acids and polypeptides are useful for diagnosing, prognosing and treating
CC infections caused by HBV genotype G. They can be used in a vaccine to
CC treat or prevent HBV genotype G infection. The HBV genotype G derived
CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
CC sample or diagnosis of HBV genotype G infection. The presence of HBV
CC genotype G statistically correlates with the presence of liver damage
CC and/or liver cancer in the subject. The HBV genotype G core insert
CC peptide encoding nucleic acid is useful for designing monitoring assays
CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
CC and HBeAg (genotype G e antigen) in patients infected with HBV. The
CC antibodies or antigens of HBV genotype G are useful for identifying a
CC stage of liver disease caused by HBV genotype G. The present sequence is
CC a hepatitis B virus (HBV) strain FRI, genotype G DNA fragment
XX
XX Sequence 129 BP; 25 A; 32 C; 26 G; 46 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 30; DB 4; Length 129;
Best Local Similarity 86.7%; Pred. No. 0.0038;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUACAAGAGAGUAGGACAGGT 30
DB 43 GACATGAACAAGAGATGATTAGGCAGGT 14

RESULT 6
AAD27422/c
ID AAD27422 standard; DNA; 639 BP.
XX
XX AAD27422;
AC
XX
XX 18-APR-2002 (first entry)
DT
XX
XX Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
DE
XX
XX Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW antiprotozoal; ds.
XX
XX Hepatitis B virus.
OS
XX
XX
XX
XX Key Location/Qualifiers
FH 1..639
CDS /*tag= a
FT /product= "HBcAg"
FT
FT sig_peptide 1..87
FT /*tag= b
FT mat_peptide 88..636
FT /*tag= c
FT /product= "Mature HBC protein"
FT
XX WO200198333-A2.
PN
XX
XX 27-DEC-2001.
PD
XX
XX 22-JUN-2001; 2001WO-GB002817.
PF
XX
XX 22-JUN-2000; 2000GB-00015308.
PR

```

```

PR 06-OCT-2000; 2000GB-00024544.
XX
XX (CELL-) CELLTECH PHARM LTD.
XX
XX Page M, Li J, Pumpens P;
PI
XX
XX WPI; 2002-098223/13.
DR
XX
XX P-PSDB; AAE17018.
DR
XX
XX New proteins comprising a modified hepatitis B core antigen, useful as a
PT vaccine in prophylactic or therapeutic vaccination of the human or animal
PT body, particularly against hepatitis B virus infection.
XX
XX Disclosure; Page 38-39; 40pp; English.
PS
XX
XX The invention relates to modified proteins comprising hepatitis B virus
CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC repeats has been deleted and the protein comprising the C-terminal
CC cysteine of HBcAg. The deleted region may be replaced by an epitope from
CC a protein other than HBcAg, in which case the HBcAg acts as a carrier to
CC present the epitope to the immune system. This chimeric protein or its
CC nucleic acid is useful as a vaccine or in a method of prophylactic or
CC therapeutic vaccination of the human or animal body, particularly against
CC HBV. The nucleic acid encoding the protein may be used in gene therapy or
CC DNA vaccination protocols. The chimeric protein or its nucleic acid may
CC also be used as the basis of a prophylactic vaccine against a range of
CC diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
CC influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
CC immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
CC tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
CC diarrhoea, meningitis or gonorrhea. The present sequence is a DNA
CC encoding Hepatitis B virus core antigen (HBcAg)
XX
XX Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 30; DB 6; Length 639;
Best Local Similarity 86.7%; Pred. No. 0.0049;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGACACAGAGATGATTAGGCAGGT 30
DB 43 GACATGACACAGAGATGATTAGGCAGGT 14

RESULT 7
AAD31509/c
ID AAD31509 standard; DNA; 639 BP.
XX
XX AAD31509;
AC
XX
XX 18-JUN-2002 (first entry)
DT
XX
XX Hepatitis B virus core antigen (HBcAg) encoding DNA.
DE
XX
XX Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
KW therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
KW polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX
XX Hepatitis B virus.
OS
XX
XX
XX
XX Key Location/Qualifiers
FH 1..639
CDS /*tag= a
FT /product= "Hbc protein"
FT
FT sig_peptide 1..87
FT /*tag= b
FT mat_peptide 88..636
FT /*tag= c
FT /product= "Mature HBC protein"
FT
XX WO200177158-A1.
PN
XX
XX 18-OCT-2001.
PR

```

XX	09-APR-2001; 2001WO-GB001607.
PF	(MEDE-) MEDEVA EURO LTD.
XX	Gehin A, Gilbert R, Stuart D, Rowlands D;
XX	P-PSDB; AAEL9793.
XX	Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
PT	prophylactic or therapeutic treatment of humans or animals against e.g.
PT	HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
PT	mouth disease.
XX	Disclosure; Page 23-24; 27pp; English.
XX	The present invention relates to hepatitis B virus (HBV) core antigen
CC	(HbcAg) fusion proteins and polynucleotides encoding such proteins.
CC	Sequences of the invention are useful in methods of prophylactic or
CC	therapeutic vaccination or to manufacture medicaments for prophylactic or
CC	therapeutic vaccination of the human or animal body against HBV, e.g.
CC	against viral hepatitis. They are also useful as a prophylactic vaccine
CC	against e.g. hepatitis c virus, influenza, polio, herpes, rabies,
CC	acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
CC	present sequence is a DNA encoding hepatitis B virus core antigen (HbcAg)
XX	
SQ	Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;
	Query Match 100.0%; Score 30; DB 6; Length 639;
	Best Local Similarity 86.7%; Pred. No. 0.0049;
	Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GACAUGAACAGAGAUAUAGGCAGAGGT 30
Dd	: :: :
	43 GACATGACACAGAGATGATTAGGCAGAGT 14
	RESULT 8
AHH77569/c	
ID	AAH77569 standard; DNA; 655 BP.
XX	
AC	AAH77569;
XX	
DT	19-OCT-2001 (first entry)
XX	
DE	HBV genotype G strain US1 preCore/Core DNA.
XX	
KW	Hepatitis B virus; HBV; preCore; Core; pres1; pres2; HBS; HBx; HBPol;
KW	HBSAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBCAg;
KW	HBeAg; ds.
XX	
OS	Hepatitis B virus.
XX	
PN	WO200140279-A2.
XX	
PD	07-JUN-2001.
XX	
PF	20-NOV-2000; 2000WO-EP011526.
XX	
PR	03-DEC-1999; 99EP-00870252.
PR	07-DEC-1999; 99US-0169287P.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Stuyver L, Van Geyt C, De Gendt S;
XX	
DR	WI; 2001-374785/39.
XX	
PT	Novel isolated and/or purified hepatitis B virus polypeptide and
PT	polynucleotide sequences that are phylogenetically different from HBV
XX	genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
XX	therapy.
XX	Claim 3; Fig 7; 9app; English.
XX	
XX	The invention relates to the complete nucleic acid sequence of a new
CC	human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC	This genotype was found with a high prevalence in patients chronically
CC	infected with HBV and residing in Europe and the USA. The invention
CC	relates to a fully defined sequence of 3248 nucleotides as given in
CC	specification, a sequence with 92% identity to the given sequence, or
CC	sequence that is degenerate to the mentioned sequences. These
CC	polynucleotides are useful for HBV genotyping. The proteins encoded by
CC	the polynucleotides are useful for detecting antibodies in a biological
CC	sample. Ligands that bind to the proteins and antibodies directed against
CC	the proteins are useful for detecting the proteins and for detecting
CC	HBCAg and HBeAg (precursor proteins). They are also useful for
CC	preparing a vaccine or medication for treating HBV infections. The
CC	present sequence is provided in an alignment of preCore/Core sequences of
CC	an HBV genotype A strain (HBVXCP5) and 7 strains (FR1, FR2, US1, US3,
CC	US6, US7, US9, US10) of HBV genotype G
XX	
SQ	Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;
	Query Match 100.0%; Score 30; DB 4; Length 655;
	Best Local Similarity 86.7%; Pred. No. 0.005;
	Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GACAUGAACAGAGAUAUAGGCAGAGGT 30
Dd	: :: :
	43 GACATGACACAGAGATGATTAGGCAGAGT 14
	RESULT 9
AHH77568/c	
ID	AAH77568 standard; DNA; 655 BP.
XX	
AC	AAH77568;
XX	
DT	19-OCT-2001 (first entry)
XX	
DE	HBV genotype G strain FR2 preCore/Core DNA.
XX	
KW	Hepatitis B virus; HBV; preCore; Core; pres1; pres2; HBS; HBx; HBPol;
KW	HBSAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBCAg;
KW	HBeAg; ds.
XX	
OS	Hepatitis B virus.
XX	
PN	WO200140279-A2.
XX	
PD	07-JUN-2001.
XX	
PF	20-NOV-2000; 2000WO-EP011526.
XX	
PR	03-DEC-1999; 99EP-00870252.
PR	07-DEC-1999; 99US-0169287P.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Stuyver L, Van Geyt C, De Gendt S;
XX	
DR	WI; 2001-374785/39.
XX	
PT	Novel isolated and/or purified hepatitis B virus polypeptide and
PT	polynucleotide sequences that are phylogenetically different from HBV
XX	genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
XX	therapy.
XX	Claim 3; Fig 7; 9app; English.
XX	
XX	The invention relates to the complete nucleic acid sequence of a new
CC	human hepatitis B virus (HBV) genotype, provisionally named genotype G.
CC	This genotype was found with a high prevalence in patients chronically
CC	infected with HBV and residing in Europe and the USA. The invention
CC	relates to a fully defined sequence of 3248 nucleotides as given in
CC	specification, a sequence with 92% identity to the given sequence, or
CC	sequence that is degenerate to the mentioned sequences. These
CC	polynucleotides are useful for HBV genotyping. The proteins encoded by
CC	the polynucleotides are useful for detecting antibodies in a biological
CC	sample. Ligands that bind to the proteins and antibodies directed against
CC	the proteins are useful for detecting the proteins and for detecting
CC	HBCAg and HBeAg (precursor proteins). They are also useful for
CC	preparing a vaccine or medication for treating HBV infections. The
CC	present sequence is provided in an alignment of preCore/Core sequences of
CC	an HBV genotype A strain (HBVXCP5) and 7 strains (FR1, FR2, US1, US3,
CC	US6, US7, US9, US10) of HBV genotype G
XX	
SQ	Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;
	Query Match 100.0%; Score 30; DB 4; Length 655;
	Best Local Similarity 86.7%; Pred. No. 0.005;
	Matches 26; Conservative 4; Mismatches

CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for HBV genotyping. The proteins encoded by
 CC the polynucleotides are useful for detecting antibodies in a biological
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBcAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is provided in an alignment of preCore/Core sequences of
 CC an HBV genotype A strain (HBVXCPs) and 7 strains (FR1, FR2, US1, US3,
 CC US6, US7, US9, US10) of HBV genotype G

SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 30; DB 4; Length 655;
 Best Local Similarity 86.7%; Pred. No. 0.005;
 Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGCAUGAUUAGGCGAGGT 30
 |||||:|||||:|||||:|||||:|||||
 Db 43 GACATGAACAAGATGATTAGGCGAGGT 14

RESULT 10

AAH77574/c
 ID AAH77574 standard; DNA; 655 BP.

AC AAH77574;

XX 19-OCT-2001 (first entry)

XX HBV genotype G strain US10 preCore/Core DNA.

KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPOL;
 KW HBSAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
 KW HBeAg; ds.

XX Hepatitis B virus.

XX WO200140279-A2.

XX 07-JUN-2001.

XX 20-NOV-2000; 2000WO-EP011526.

XX 03-DEC-1999; 99EP-00870252.

XX 07-DEC-1999; 99US-0169287P.

XX (INNO-) INNOGENETICS NV.

XX Stuyver L, Van Geyt C, De Gendt S;

XX WPI; 2001-374785/39.

XX Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.

XX Claim 3; Fig 7; 94pp; English.

XX The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
 CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for HBV genotyping. The proteins encoded by
 CC the polynucleotides are useful for detecting antibodies in a biological

CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBcAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is provided in an alignment of preCore/Core sequences of
 CC an HBV genotype A strain (HBVXCPs) and 7 strains (FR1, FR2, US1, US3,
 CC US6, US7, US9, US10) of HBV genotype G

SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;

Query Match 100.0%; Score 30; DB 4; Length 655;
 Best Local Similarity 86.7%; Pred. No. 0.005;
 Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGCAUGAUUAGGCGAGGT 30
 |||||:|||||:|||||:|||||:|||||
 Db 43 GACATGAACAAGATGATTAGGCGAGGT 14

RESULT 11

AAH77573/c

ID AAH77573 standard; DNA; 655 BP.

XX AC AAH77573;

XX 19-OCT-2001 (first entry)

XX HBV genotype G strain US7 preCore/Core DNA.

KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPOL;
 KW HBSAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
 KW HBeAg; ds.

XX Hepatitis B virus.

XX WO200140279-A2.

XX 07-JUN-2001.

XX 20-NOV-2000; 2000WO-EP011526.

XX 03-DEC-1999; 99EP-00870252.

XX 07-DEC-1999; 99US-0169287P.

XX (INNO-) INNOGENETICS NV.

XX Stuyver L, Van Geyt C, De Gendt S;

XX WPI; 2001-374785/39.

XX Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.

XX Claim 3; Fig 7; 94pp; English.

XX The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
 CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for HBV genotyping. The proteins encoded by
 CC the polynucleotides are useful for detecting antibodies in a biological
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBcAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is provided in an alignment of preCore/Core sequences of
 CC an HBV genotype A strain (HBVXCPs) and 7 strains (FR1, FR2, US1, US3,
 CC US6, US7, US9, US10) of HBV genotype G

```
XX SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;
Query Match 100.0%; Score 30; DB 4; Length 655;
Best Local Similarity 86.7%; Pred. No. 0.005;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAGAGAGUAGUAGGCAGAGGT 30
Db 43 GACATGAACAAGAGATGATTAGGCAGAGGT 14

RESULT 12
AAH77570/c
ID AAH77570 standard; DNA; 655 BP.
XX AC AAH77570;
XX DT 19-OCT-2001 (first entry)
XX HBV genotype G strain US3 preCore/Core DNA.
XX Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
XX HBeAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBeAg;
XX Hepatitis B virus.
XX WO200140279-A2.
XX 07-JUN-2001.
XX 20-NOV-2000; 2000WO-EP011526.
XX 03-DEC-1999; 99EP-00870252.
XX 07-DEC-1999; 99US-0169287P.
XX (INNO-) INNOGENETICS NV.
XX Stuyver L, Van Geyt C, De Gendt S;
XX WPI; 2001-374785/39.
XX Novel isolated and/or purified hepatitis B virus polypeptide and
XX polynucleotide sequences that are phylogenetically different from HBV
XX genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
XX therapy.
XX Claim 3; Fig 7; 94pp; English.
XX The invention relates to the complete nucleic acid sequence of a new
XX human hepatitis B virus (HBV) genotype, provisionally named genotype G.
XX This genotype was found with a high prevalence in patients chronically
XX infected with HBV and residing in Europe and the USA. The invention
XX relates to a fully defined sequence of 3248 nucleotides as given in
XX specification, a sequence with 92% identity to the given sequence, or
XX polynucleotides are useful for HBV genotyping. The proteins encoded by
XX the polynucleotides are useful for detecting antibodies in a biological
XX sample. Ligands that bind to the proteins and antibodies directed against
XX HBeAg and HBeAg (precore precursor proteins). They are also useful for
XX preparing a vaccine or medicament for treating HBV infections. The
XX present sequence is provided in an alignment of preCore/Core sequences of
XX an HBV genotype A strain (HBVXCP8) and 7 strains (FR1, FR2, US1, US3,
XX US6, US7, US9, US10) of HBV genotype G
XX SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;
Query Match 100.0%; Score 30; DB 4; Length 655;
Best Local Similarity 86.7%; Pred. No. 0.005;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 GACAUGAACAGAGAGUAGUAGGCAGAGGT 30
Db 43 GACATGAACAAGAGATGATTAGGCAGAGGT 14

RESULT 13
AAH77571/c
ID AAH77571 standard; DNA; 655 BP.
XX AC AAH77571;
XX DT 19-OCT-2001 (first entry)
XX HBV genotype G strain US5 preCore/Core DNA.
XX Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBPol;
XX HBeAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBeAg;
XX Hepatitis B virus.
XX WO200140279-A2.
XX 07-JUN-2001.
XX 20-NOV-2000; 2000WO-EP011526.
XX 03-DEC-1999; 99EP-00870252.
XX 07-DEC-1999; 99US-0169287P.
XX (INNO-) INNOGENETICS NV.
XX Stuyver L, Van Geyt C, De Gendt S;
XX WPI; 2001-374785/39.
XX Novel isolated and/or purified hepatitis B virus polypeptide and
XX polynucleotide sequences that are phylogenetically different from HBV
XX genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
XX therapy.
XX Claim 3; Fig 7; 94pp; English.
XX The invention relates to the complete nucleic acid sequence of a new
XX human hepatitis B virus (HBV) genotype, provisionally named genotype G.
XX This genotype was found with a high prevalence in patients chronically
XX infected with HBV and residing in Europe and the USA. The invention
XX relates to a fully defined sequence of 3248 nucleotides as given in
XX specification, a sequence with 92% identity to the given sequence, or
XX polynucleotides are useful for HBV genotyping. The proteins encoded by
XX the polynucleotides are useful for detecting antibodies in a biological
XX sample. Ligands that bind to the proteins and antibodies directed against
XX HBeAg and HBeAg (precore precursor proteins). They are also useful for
XX preparing a vaccine or medicament for treating HBV infections. The
XX present sequence is provided in an alignment of preCore/Core sequences of
XX an HBV genotype A strain (HBVXCP8) and 7 strains (FR1, FR2, US1, US3,
XX US6, US7, US9, US10) of HBV genotype G
XX SQ Sequence 655 BP; 138 A; 154 C; 140 G; 195 T; 0 U; 28 Other;
Query Match 100.0%; Score 30; DB 4; Length 655;
Best Local Similarity 86.7%; Pred. No. 0.005;
Matches 26; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAUGAACAGAGAGUAGUAGGCAGAGGT 30
Db 43 GACATGAACAAGAGATGATTAGGCAGAGGT 14

RESULT 14
AAH7572/c
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-60

Perfect score: 20
Sequence: 1 gacatgaacaagagatgatt 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	20	100.0	30	6	AR027810	AR027810 Sequence
2	20	100.0	30	6	AR027841	AR027841 Sequence
3	20	100.0	87	6	AX151115	AX151115 Sequence
4	20	100.0	99	14	HPBRSCA	M76687 Hepatitis B
5	20	100.0	99	14	HPBRSCB	M76688 Hepatitis B
6	20	100.0	99	14	HPBRSCC	M76689 Hepatitis B
7	20	100.0	99	14	HPBRSCD	M76690 Hepatitis B
8	20	100.0	99	14	HPBRSECF	M76691 Hepatitis B
9	20	100.0	99	14	HPBRSECG	M76692 Hepatitis B
10	20	100.0	99	14	HPBRSECH	M76693 Hepatitis B
11	20	100.0	99	14	HPBRSECI	M76694 Hepatitis B
12	20	100.0	99	14	HPBRSECM	M76695 Hepatitis B
13	20	100.0	99	14	HPBRSECM	M76699 Hepatitis B
14	20	100.0	129	6	AX151114	AX151114 Sequence
15	20	100.0	150	14	AF528205	AF528205 Hepatitis
16	20	100.0	150	14	AF528206	AF528206 Hepatitis
17	20	100.0	150	14	AF528207	AF528207 Hepatitis
18	20	100.0	150	14	AF528208	AF528208 Hepatitis
19	20	100.0	150	14	AF528209	AF528209 Hepatitis

C 20	20	100.0	150	14	AF528210	Hepatitis
C 21	20	100.0	150	14	AF528211	Hepatitis
C 22	20	100.0	150	14	AF528212	Hepatitis
C 23	20	100.0	150	14	AF528213	Hepatitis
C 24	20	100.0	150	14	AF528214	Hepatitis
C 25	20	100.0	150	14	AF528215	Hepatitis
C 26	20	100.0	150	14	AF528216	Hepatitis
C 27	20	100.0	150	14	AF528217	Hepatitis
C 28	20	100.0	150	14	AF528218	Hepatitis
C 29	20	100.0	150	14	AF528219	Hepatitis
C 30	20	100.0	150	14	AF528220	Hepatitis
C 31	20	100.0	150	14	AF528221	Hepatitis
C 32	20	100.0	150	14	AF528222	Hepatitis
C 33	20	100.0	150	14	AF528224	Hepatitis
C 34	20	100.0	150	14	AF528225	Hepatitis
C 35	20	100.0	150	14	AF528226	Hepatitis
C 36	20	100.0	150	14	AF528227	Hepatitis
C 37	20	100.0	150	14	AF528228	Hepatitis
C 38	20	100.0	150	14	AF528229	Hepatitis
C 39	20	100.0	150	14	AF528231	Hepatitis
C 40	20	100.0	150	14	AF528232	Hepatitis
C 41	20	100.0	150	14	AF528233	Hepatitis
C 42	20	100.0	150	14	AF528234	Hepatitis
C 43	20	100.0	150	14	AF528235	Hepatitis
C 44	20	100.0	150	14	AF528236	Hepatitis
C 45	20	100.0	150	14	AF528237	Hepatitis

ALIGNMENTS

RESULT 1
AR027810
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN

AR027810
Sequence 8 from patent US 5856459.
AR027810
AR027810.1 GI:5938630
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 30)
Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and
Mills,J.S.
Oligonucleotides specific for hepatitis B virus
Patent: US 5856459-A 8 05-JAN-1999;
Location/Qualifiers
1..30
/organism="unknown"
/mol_type="unassigned DNA"

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 30;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
|||||
DB 1 GACATGAACAAGAGATGATT 20

RESULT 2
AR027841
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AR027841
Sequence 39 from patent US 5856459.
AR027841
AR027841.1 GI:5938661
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 30)
Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

Mills,J.S.
 TITLE Oligonucleotides specific for hepatitis B virus
 JOURNAL Patent: US 5856459-A 39 05-JAN-1999;
 FEATURES Location/Qualifiers
 source 1..30
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 30;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGAACAAGAGATGATT 20
 |||||
 Db 1 GACATGAACAAGAGATGATT 20

RESULT 3
 AX151115/c
 LOCUS AX151115 87 bp DNA linear PAT 22-JUN-2001
 DEFINITION Sequence 4 from Patent WO0138498.
 ACCESSION AX151115
 VERSION AX151115.1 GI:14533317
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM other sequences; artificial sequences.

REFERENCE 1
 AUTHORS Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,
 Fried,M. and Rousau,R.
 TITLE A new genotype of hepatitis b virus
 JOURNAL Patent: WO 0138498-A 4 31-MAY-2001;
 Pharmasset, Inc. (US); INNOGENETICS N.V. (BE)
 FEATURES Location/Qualifiers
 source 1..87
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 87;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGAACAAGAGATGATT 20
 |||||
 Db 43 GACATGAACAAGAGATGATT 24

RESULT 4
 HBPBPRECA/c
 LOCUS HBPBPRECA 99 bp DNA linear VRL 11-MAY-1994
 DEFINITION Hepatitis B virus typel precore protein (pre-C region, C) gene, 5'
 end.
 ACCESSION M76687
 VERSION M76687.1 GI:485341
 KEYWORDS e antigen; precore protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 REFERENCE 1 (bases 1 to 99)
 AUTHORS Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
 Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive
 carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582
 COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES Location/Qualifiers
 source 1..99
 /organism="Hepatitis B virus"

gene /mol_type="genomic DNA"
 /db_xref="taxon:10407"
 10..93
 /gene="C"
 CDS 10..93
 /gene="C"
 /standard_name="pre-C region"
 /codon_start=1
 /product="precure protein"
 /protein_id="AAA45507.1"
 /db_xref="GI:485342"
 /translation="MQLFHLCLIISCPTVQASKLCLGWL"
 92
 /gene="C"
 /note="g in wt; a in virus type 1 (creates internal stop
 codon)"

variation

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 99;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGAACAAGAGATGATT 20
 |||||
 Db 52 GACATGAACAAGAGATGATT 33

RESULT 5

HBPBPRECB/c
 LOCUS HBPBPRECB 99 bp DNA linear VRL 11-MAY-1994
 DEFINITION Hepatitis B virus type 2precure protein (pre-C region, C) gene, 5'
 end.
 ACCESSION M76688
 VERSION M76688.1 GI:485343
 KEYWORDS e antigen; precore protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 REFERENCE 1 (bases 1 to 99)
 AUTHORS Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
 Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive
 carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582
 COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES Location/Qualifiers
 source 1..99
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10407"

variation

gene

CDS

variation

gene

CDS

variation

gene

CDS

variation

gene

CDS

variation

gene

CDS

variation

gene

CDS

variation

gene

CDS

variation

QY 1 GACATGAACAAGAGATGATT 20
 |||||
 Db 52 GACATGAACAAGAGATGATT 33

RESULT 6
 HPBP/CC/c
 LOCUS
 DEFINITION Hepatitis B virus type 3precure protein (pre-C region, C) gene, 5'
 end.
 ACCESSION M76689
 VERSION M76690.1 GI:485345
 KEYWORDS e antigen; precure protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 REFERENCE 1 (bases 1 to 99)
 AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582
 COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES
 source
 Location/Qualifiers
 1..99
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10407"
 6
 /notes="c in wt; t in virus type 3"
 10..93
 /gene="C"
 10..93
 /genes="C"
 /standard_name="pre-C region"
 /codon_start=1
 /product="precure protein"
 /protein_id="AAA45509.1"
 /db_xref="GI:485346"
 /translation="MQLFHLIIISCSPTFQASKLCLGLW"
 58
 /genes="C"
 /notes="g in wt; t in virus type 3 (val to phe)"
 92
 /gene="C"
 /notes="g in wt; a in virus type 3 (creates internal stop codon)"
 variation
 variation
 Query Match 100.0%; Score 20; DB 14; Length 99;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
 |||||
 Db 52 GACATGAACAAGAGATGATT 33

RESULT 7
 HPBP/CC/c
 LOCUS
 DEFINITION Hepatitis B virus type 4 precure protein (pre-C region, C) gene, 5'
 end.
 ACCESSION M76690
 VERSION M76690.1 GI:485347
 KEYWORDS e antigen; precure protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 REFERENCE 1 (bases 1 to 99)

AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582
 COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES
 source
 Location/Qualifiers
 1..99
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10407"
 10..93
 /gene="C"
 10..93
 /genes="C"
 /standard_name="pre-C region"
 /codon_start=1
 /product="precure protein"
 /protein_id="AAA45510.1"
 /db_xref="GI:485348"
 /translation="MQLFHLIIISCSPTVQASKLCLGLW"
 92
 /genes="C"
 /notes="g in wt; a in virus type 4 (creates internal stop codon)"
 95
 /notes="g in wt; a in virus type 4 (gly to asp)"
 variation
 variation
 Query Match 100.0%; Score 20; DB 14; Length 99;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
 |||||
 Db 52 GACATGAACAAGAGATGATT 33

RESULT 8
 HPBP/CC/c
 LOCUS
 DEFINITION Hepatitis B virus type 5 precure protein (pre-C region, C) gene, 5'
 end.
 ACCESSION M76691
 VERSION M76691.1 GI:485349
 KEYWORDS e antigen; precure protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
 REFERENCE 1 (bases 1 to 99)
 AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582
 COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES
 source
 Location/Qualifiers
 1..99
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10407"
 10..93
 /gene="C"
 10..93
 /genes="C"
 /standard_name="pre-C region"
 /codon_start=1
 /product="precure protein"
 /protein_id="AAA45511.1"
 /db_xref="GI:485349"

```

/db_xref="GI:485350"
/translation="MQLFHLCLIIISCSPTVPQSKLCIGWL"
64..67
/gene="C"
/notes="gcc in wt; ccg in virus type 5 (ala to pro)"
92
/gene="C"
/notes="g in wt; a in virus type 5 (creates internal stop codon)"
95
ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GACATGAACACAGAGATGATT 20
    |||||
Db   52 GACATGAACACAGAGATGATT 33

RESULT 9
HPBPREFC/c
LOCUS      Hepatitis B virus type 6 precore protein (pre-C region, C) gene, 5'
DEFINITION
end.
ACCESSION  M76692
VERSION     M76692.1 GI:485351
KEYWORDS   e antigen; precore protein; tolerogen.
SOURCE     Hepatitis B virus
ORGANISM   Hepatitis B virus
REFERENCE  1 (bases 1 to 99)
AUTHORS    Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE      Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL    Virology 183 (2), 840-844 (1991)
MEDLINE    91306476
PUBMED     1853582
COMMENT    Original source text: Hepatitis B virus DNA.
FEATURES   Location/Qualifiers
            source
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            10..93
            /gene="C"
            10..93
            misc_feature
            10..93
            /gene="C"
            /product="precore protein"
            /standard_name="pre-C region note: putative CDS"
            10
            variation
            14
            /gene="C"
            /note="a in wt; t in virus type 7 (loss of start codon)"
            14
            /gene="C"
            /note="a in wt; g in virus type 7 (gln to arg)"
            92
            /gene="C"
            /note="g in wt; a in virus type 7 (creates internal stop codon)"
            95
ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GACATGAACACAGAGATGATT 20
    |||||
Db   52 GACATGAACACAGAGATGATT 33

RESULT 11
HPBPREFC/c
LOCUS      Hepatitis B virus type 8 precore protein (pre-C region, C) gene, 5'
DEFINITION
end.
ACCESSION  M76694
VERSION     M76694.1 GI:485353
KEYWORDS   e antigen; precore protein; tolerogen.
SOURCE     Hepatitis B virus
ORGANISM   Hepatitis B virus
REFERENCE  1 (bases 1 to 99)
AUTHORS    Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE      Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL    Virology 183 (2), 840-844 (1991)
MEDLINE    91306476
PUBMED     1853582
COMMENT    Original source text: Hepatitis B virus DNA.
FEATURES   Location/Qualifiers
            source
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            10..99
            /gene="C"
            10..99
            misc_feature
            10..99
            /gene="C"
            /product="precore protein"
            /notes="putative cds"
            11
            variation
            11
            /gene="C"
            /note="t in wt; c in virus type 6 (loss of start codon)"
            11
ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GACATGAACACAGAGATGATT 20
    |||||
Db   52 GACATGAACACAGAGATGATT 33

RESULT 10
HPBPREFC/c
LOCUS      Hepatitis B virus type 7 precore protein (pre-C region, C) gene, 5'
DEFINITION
end.
ACCESSION  M76693
VERSION     M76693.1 GI:485352

```

```

e antigen; precore protein; tolerogen.
Hepatitis B virus
Hepatitis B virus
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE  1 (bases 1 to 99)
AUTHORS    Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE      Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL    Virology 183 (2), 840-844 (1991)
MEDLINE    91306476
PUBMED     1853582
COMMENT    Original source text: Hepatitis B virus DNA.
FEATURES   Location/Qualifiers
            source
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            10..93
            /gene="C"
            10..93
            misc_feature
            10..93
            /gene="C"
            /product="precore protein"
            /standard_name="pre-C region note: putative CDS"
            10
            variation
            14
            /gene="C"
            /note="a in wt; t in virus type 7 (loss of start codon)"
            14
            /gene="C"
            /note="a in wt; g in virus type 7 (gln to arg)"
            92
            /gene="C"
            /note="g in wt; a in virus type 7 (creates internal stop codon)"
            95
ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GACATGAACACAGAGATGATT 20
    |||||
Db   52 GACATGAACACAGAGATGATT 33

RESULT 11
HPBPREFC/c
LOCUS      Hepatitis B virus type 8 precore protein (pre-C region, C) gene, 5'
DEFINITION
end.
ACCESSION  M76694
VERSION     M76694.1 GI:485353
KEYWORDS   e antigen; precore protein; tolerogen.
SOURCE     Hepatitis B virus
ORGANISM   Hepatitis B virus
REFERENCE  1 (bases 1 to 99)
AUTHORS    Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE      Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL    Virology 183 (2), 840-844 (1991)
MEDLINE    91306476
PUBMED     1853582
COMMENT    Original source text: Hepatitis B virus DNA.
FEATURES   Location/Qualifiers
            source
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            10..93
            /gene="C"
            10..93
            misc_feature
            10..93

```

```
/gene="C"
/product="precure protein"
/standard_name="pre-C region note: putative CDS"
12
variation
/gene="C"
/notes="g in wt; t in virus type 8 (loss of start codon)"
92
variation
/gene="C"
/notes="g in wt; a in virus type 8 (creates internal stop codon)"
95
variation
/notes="g in wt; a in virus type 8 (gly to asp)"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATGAACAAGAGATGATT 20
|||||
Db 52 GACATGAACAAGAGATGATT 33

RESULT 12
HPBPREC1/c
LOCUS
DEFINITION
Hepatitis B virus type 9 precure protein (pre-C region, C) gene, 5' end.
ACCESSION M76695.1 GI:485354
VERSION
KEYWORDS e antigen; precure protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
source
Location/Qualifiers
1..99
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:10407"
10..99
/gene="C"
10..>99
/genes="C"
/standard_name="pre-C region"
/codon_start=1
/product="precure protein"
/protein_id="AAA4515.1"
/db_xref="GI:485362"
/translation="MOLFHLIIISCSCTVQASKLGLWLDWM"
95
variation
/gene="C"
/notes="g in wt; a in virus type 13 (gly to asp)"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATGAACAAGAGATGATT 20
|||||
Db 52 GACATGAACAAGAGATGATT 33

misc_feature
/gene="C"
10..93
/genes="C"
/standard_name="precure protein"
/standard_name="pre-C region note: putative CDS"
13
variation
/gene="C"
/notes="c in wt; t in virus type 9 (creates internal stop codon)"
92
variation
/gene="C"
/notes="g in wt; a in virus type 9 (creates internal stop codon)"
95
variation
/notes="g in wt; a in virus type 9 (gly to asp)"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATGAACAAGAGATGATT 20
|||||
Db 52 GACATGAACAAGAGATGATT 33
```

```
RESULT 13
HPBPREC1/c
LOCUS
DEFINITION
Hepatitis B virus type 13 precure protein (pre-C region, C) gene, 5' end.
ACCESSION M76699
VERSION M76699.1 GI:485361
KEYWORDS e antigen; precure protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
source
Location/Qualifiers
1..99
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:10407"
10..99
/gene="C"
10..>99
/genes="C"
/standard_name="pre-C region"
/codon_start=1
/product="precure protein"
/protein_id="AAA4515.1"
/db_xref="GI:485362"
/translation="MOLFHLIIISCSCTVQASKLGLWLDWM"
95
variation
/gene="C"
/notes="g in wt; a in virus type 13 (gly to asp)"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATGAACAAGAGATGATT 20
|||||
Db 52 GACATGAACAAGAGATGATT 33

RESULT 14
AX151114/c
LOCUS
DEFINITION
Sequence 3 from Patent WO0138498.
ACCESSION AX151114
VERSION AX151114.1 GI:14533316
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F., Fried,M. and Rossau,R.
TITLE A new genotype of hepatitis b virus
JOURNAL Patent: WO 0138498-A 3 31-MAY-2001; Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES
source
Location/Qualifiers
1..129
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN
```

```

Query Match      100.0%; Score 20; DB 6; Length 129;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GACATGAACAAGAGATGATT 20
      |||
Db      43 GACATGAACAAGAGATGATT 24

RESULT 15
AF528205/c
LOCUS      AF528205      150 bp      DNA      linear      VRL 31-JUL-2003
DEFINITION Hepatitis B virus ASC1123 core antigen precursor, gene, partial
            cds.
ACCESSION  AF528205
VERSION    AF528205.1 GI:32810971
KEYWORDS   Hepatitis B virus
SOURCE     Hepatitis B virus
ORGANISM   Hepatitis B virus
REFERENCE  1 (bases 1 to 150)
            Viruses; Retroviridae; Hepadnaviridae; Orthohepadnavirus.
AUTHORS   Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
TITLE     Comparative evaluation of HBV precore and basal core promoter
            mutants in Indian patients with diverse clinical manifestations
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 150)
            Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
AUTHORS   Direct Submission
TITLE     Submitted (11-JUL-2002) Hepatitis Division, National Institute of
            Virology, 20-A, Dr Ambedkar Road, Pune, Maharashtra 411001, India
FEATURES   source
            1..150
            /organism="Hepatitis B virus"
            /proviral
            /mol_type="genomic DNA"
            /isolate="ASC1123"
            /isolation_source="asymptomatic HBsAg carrier"
            /specific_host="Homo sapiens"
            /db_xref="taxon:10407"
            /country="India"
            <1..>150
            /note="contains partial basal core promoter"
            64..>150
            /note="contains complete precore region"
            /codon_start=1
            /product="core antigen precursor"
            /protein_id="AAP87556.1"
            /db_xref="GI:32810972"
            /translation="MQLFHLCLIISCSPTVQASKLCGLXG"

ORIGIN

Query Match      100.0%; Score 20; DB 14; Length 150;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GACATGAACAAGAGATGATT 20
      |||
Db      106 GACATGAACAAGAGATGATT 87

```

Search completed: March 17, 2005, 08:14:16
 Job time : 684.733 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 2079.4 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-59
Perfect score: 30
Sequence: 1 gacaaagacagaagaauagcgaggt 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	73.3	661	9	CG452180 OG5DV84TC
C 2	22	73.3	663	6	CD304890 StrPu691.
C 3	22	73.3	701	8	AQ993161 RPCI-23-3
C 4	22	73.3	888	9	CG359992 OGLX42TV
C 5	21.2	70.7	577	4	BI378081 BPLG3_001
C 6	21.2	70.7	843	9	CW009369 ZMMB001
C 7	21	70.0	214	2	AW342249 GchEst1.G
C 8	21	70.0	345	5	BY143963 BY143963
C 9	21	70.0	399	6	CD296187 StrPu691.
C 10	21	70.0	519	2	BF103720 601647304
C 11	21	70.0	633	9	CR048521 Forward s
C 12	21	70.0	880	4	BGI04640 602311331
C 13	21	70.0	1110	8	CC258547 CH261-63K
C 14	21	70.0	1154	2	BF342749 602015013
C 15	20.8	69.3	73	9	CG623002 OST323776
C 16	20.8	69.3	715	8	AQ939213 NL1-120R
C 17	20.6	68.7	223	9	CE129305 tigr-gss-
C 18	20.6	68.7	370	9	CL423226 RP11-413M
C 19	20.6	68.7	558	8	AQ272786 nbxb0028N
C 20	20.6	68.7	577	6	CA044022 ssalpa50
C 21	20.6	68.7	595	8	AQ509958 nbxb0094H
C 22	20.6	68.7	600	8	AQ257439 nbxb0018A
C 23	20.6	68.7	622	9	CL718440 OB_BBa004
C 24	20.6	68.7	655	6	CA061518 ssalrgb52

C 25	20.6	68.7	663	6	CA061593
C 26	20.6	68.7	680	9	CE792348
C 27	20.6	68.7	699	9	CL779656
C 28	20.6	68.7	705	1	CL612489
C 29	20.6	68.7	715	1	AV707321
C 30	20.6	68.7	733	6	CA348548
C 31	20.6	68.7	757	6	CB512505
C 32	20.6	68.7	767	9	CL787302
C 33	20.6	68.7	776	9	CL793261
C 34	20.6	68.7	780	8	AQ687676
C 35	20.6	68.7	854	9	CL795294
C 36	20.6	68.7	1012	7	CN247122
C 37	20.4	68.0	344	8	BZ798022
C 38	20.4	68.0	440	8	BZ642588
C 39	20.4	68.0	490	6	CD339264
C 40	20.4	68.0	501	6	CD330453
C 41	20.4	68.0	510	1	AA003912
C 42	20.4	68.0	572	6	CD309830
C 43	20.4	68.0	586	9	FR0003872
C 44	20.4	68.0	678	6	CD310469
C 45	20.4	68.0	727	9	CL557140

ALIGNMENTS

RESULT 1
CG452180/c
LOCUS CG452180 661 bp DNA linear GSS 17-SEP-2003
DEFINITION OG5DV84TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMB0841N24,
genomic survey sequence.
ACCESSION CG452180
VERSION CG452180.1 GI:34837180
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 661)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source
1..661
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMB0841N24"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 73.3%; Score 22; DB 9; Length 661;
Best Local Similarity 76.7%; Pred. No. 1.6e+02;
Matches 23; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGAUAGGACAGGT 30
DB 481 GTCTGAGAGAGATGAGAGACAGGT 452

```

RESULT 2
CD304890
LOCUS
DEFINITION
StrPu691.001255 Sea urchin larva cDNA library MPMPp691
Strongylocentrotus purpuratus cDNA clone
MPMPp691C0520;MPI_SURUDI_20C5 5', mRNA sequence.
CD304890
CD304890.1 GI:34749939
EST.
Strongylocentrotus purpuratus
Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 663)
Poustka A.J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A.,
Reinhardt, R., Herwig, R., Fanpoulou, G. and Lehrach, H.
Generation, annotation, evolutionary analysis, and database
integration of 20,000 unique sea urchin EST clusters
Genome Res. 13 (12), 2736-2746 (2003)
Contact: Poustka AJ
laboraty 145, dept Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inestr. 63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: poustka@molgen.mpg.de
The library was characterized by oligonucleotide fingerprinting
(ONF) to reduce sequencing redundancy. According to the ONF
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONF cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of
the other clones assigned to the same ONF cluster as the clone from
which the above EST is generated is available at the sea urchin
project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA
clones and filters are distributed via the Resource Center/Primary
Database of the German Human Genome Project (http://www.rzpd.de)
PCR Primers
FORWARD: 5' CCCAGGCTTTACACTTATGCTTCGGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACCCAGCTGGCGAAGGGGATGTG 3' (M13RSP) 3'-seq
Seq primer: 5'-CCGGTCCGGATTCGCCGGT-3' pSport3/86
High quality sequence stop: 663.
Location/Qualifiers
1. 663
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon.7668"
/clone="MPMPp691C0520;MPI_SURUDI_20C5"
/tissue_type="whole larva"
/dev_stage="larva 2-3 weeks"
/lab_host="E.coli, XL1 blue"
/clone_lib="sea urchin larva cDNA library MPMPp691"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI; Random
primed and directionally cloned in pSport1 vector using a
NotI (5'-pGACTAGTTTATGATCGAGCGCCGCC (T)15-3' and a
SalI 5'-TCGACCAACGCGCTCG-3'adapters (Gibco BRL)"

Query Match 73.3%; Score 22; DB 6; Length 663;
Best Local Similarity 76.7%; Pred. No. 1.6e+02;
Matches 23; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUGAUUAGCGAGGT 30
|||||
DB 123 GACAGAGAGGAGAGAGATTAGGAGAGGT 152

RESULT 3
AQ993161/c
LOCUS
DEFINITION
RPCI-23-370H16.TV RPCI-23 Mus musculus genomic clone
RPCI-23-370H16, genomic survey sequence.
AQ993161
VERSION
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 701)
Zhao, S., Nieman, W., Feldblyum, T., Malek, J., Shatsman, S., de
Akinret, B., Levins, M., McGann, S., Tsagay, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSSs: RPCI-23-370H16.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac ends/mouse/bac_end_intro.html
Plate: 370 row: H column: 16
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. 701
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-370H16"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/notes="Organ: Kidney/Brain; Vector: pBACE3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methyase. Size
selected DNA was cloned into the pBACE3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

Query Match 73.3%; Score 22; DB 8; Length 701;
Best Local Similarity 73.3%; Pred. No. 1.7e+02;
Matches 22; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUGAUUAGCGAGGT 30
|||||
DB 519 GAAGTGAACAGAGAAGATTAGGGGAGGT 490

RESULT 4
CG359992
LOCUS
DEFINITION
OG1CX42TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBwa0733H12,
genomic survey sequence.
CG359992
VERSION
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

```

```

REFERENCE
AUTHORS    1 (bases 1 to 888)
            Whitlaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE      Consortium for Maize Genomics
JOURNAL    Unpublished (2002)
COMMENT    Other_GSSs: OG1CX42TH
            Contact: Cathy Whitlaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitlaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES   Location/Qualifiers
source     1..888
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBLA0733H12"
            /clone_lib="ZM 0.7 1.5 KB"
            /note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"
ORIGIN
Query Match      73.3%; Score 22; DB 9; Length 888;
Best Local Similarity 76.7%; Pred.No. 1.7e+02;
Matches 23; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGAGAUUAGGACAGAGGT 30
   ||||| ||||| ||||| ||||| |||||
Db 245 GTCATGAAGAGAGATGAAGACGAGAGGT 274

RESULT 5
BI378081/c
LOCUS      BI378081
DEFINITION 577 bp mRNA linear EST 26-AUG-2003
            BFLG3 001884 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG
            or MPMPG498) Branchiostoma floridae cDNA clone MPMPG498N0628 5',
            mRNA sequence.
ACCESSION  BI378081
VERSION    BI378081.1 GI:30913195
KEYWORDS   EST.
SOURCE     Branchiostoma floridae (Florida lancelet)
ORGANISM   Branchiostoma floridae
            Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
            Branchiostoma.
REFERENCE  1 (bases 1 to 577)
AUTHORS   Panopoulou,G., Hennig,S., Groth,D., Krause,A., Foustka,A.J.,
            Herwig,R., Vingron,M. and Lehrach,H.
TITLE     New evidence for genome-wide duplications at the origin of
            vertebrates using an amphioxus gene set and completed animal
            genomes
JOURNAL   Genome Res. 13 (6A), 1056-1066 (2003)
MEDLINE  22683279
PUBMED   12799346
COMMENT   Contact: Panopoulou G
            Laboratory 145, Dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1235
            Fax: +49 30 8413 1128
            Email: panopoul@molgen.mpg.de
            The library was characterised by oligonucleotide fingerprinting
            (ONFP) to reduce sequencing redundancy. According to the ONFP
            procedure, clones giving the same hybridisation pattern with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per cluster is selected for sequencing. The size of each
            cluster is an indicator of the frequency of a transcript in the
            analysed library. The cluster size as well the coordinates of the
            rest of the clones assigned to the same fingerprint cluster as the

```

```

clone from which the above EST is generated is available at the
amphioxus project site at: http://www.molgen.mpg.de/amphioxus/
amplifiers and filters are distributed via the Resource Center/Primary
Database of the German Genome Project (http://www.rzpd.de)
PCR Primers
FORWARD: 5' CCCAGGCTTACACTTATGCTTCGGCTCG 3' (M13RSP)
BACKWARD: 5' GCTATTACCGACGCTGCGAAGGGGATGTG 3' (M13FSP)
Insert length: 1200 Std Error: 0.00
Seq primer: 5'-CCGTCGCGAATTCGCGGT-3', pSport3/86
High quality sequence stop: 577.
FEATURES   Location/Qualifiers
source     1..577
            /organism="Branchiostoma floridae"
            /mol_type="mRNA"
            /db_xref="taxon:7739"
            /clone="MPMPG498N0628"
            /tissue_type="whole embryo"
            /dev_stage="5-6 hrs (gastrula stage)"
            /lab_host="E.coli, XLI blue"
            /clone_lib="Amphioxus 5-6 hrs cDNA library (Name
            convention: BFLG or MPMPG498)"
            /note="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5');
            Site 2: NotI, BamHI, HindIII (3'); OligodT primed and
            directionally cloned in pSport1 vector using a NotI
            (5'-pGACTAGTCTTAGATCGGAGCGGCCGCC (T)15-3' and a SalI 5'-
            TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."
```

ORIGIN

```

Query Match      70.7%; Score 21.2; DB 4; Length 577;
Best Local Similarity 73.1%; Pred.No. 3.5e+02;
Matches 19; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 4 AUGACAAGAGAGAUUAGGACAGAG 29
   ||||| ||||| ||||| |||||
Db 507 ATGAACAAGAGATGATTGTCAGAG 482

```

RESULT 6

```

CW009369
LOCUS      ZMMBLA0012A22.f ZMMBLA Zea mays genomic clone ZMMBLA0012A22 5',
DEFINITION genomic survey sequence.
ACCESSION  CW009369
VERSION    CW009369.1 GI:52591950
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 843)
AUTHORS   Wing,R., Luo,M., Soderlund,C. and Haller,K.
TITLE     ZMMBL sequences
JOURNAL   Unpublished (2004)
COMMENT   Contact: Rod A. Wing
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: http://genome.arizona.edu
            Plate: 0012 row: A column: 22
            Class: BAC ends.
FEATURES   Location/Qualifiers
source     1..843
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /cultivar="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBLA0012A22"
            /tissue_type="immature ears"
            /lab_host="DH10B T1 phage resistant"
            /clone_lib="ZMMBLA"

```


Best Local Similarity 75.9%; Pred. No. 3.9e+02; Mismatches 2; Conservative 2; Indels 0; Gaps 0;

QY 1 GACAUGAACAAAGAGAGAUUAGGACAGAGG 29
||||:|||||||:|:|||||||
Db 72 GACATGCACAAAGAGAGAAATGAGCCAGAGG 100

RESULT 9
CD296187
LOCUS
DEFINITION
StrP691.007582 Sea urchin larva cDNA library MPMGP691
Strongylocentrotus purpuratus cDNA clone
MPMGp691G19122:MPI_SURUDI_122G19 5', mRNA sequence.
CD296187
CD296187.1 GI:34747284

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.

REFERENCE
1 (bases 1 to 399)
Poustka A.J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A.,
Reinhardt, R., Herwig, R., Panopoulou, G. and Lehrach, H.
Generation, annotation, evolutionary analysis, and database
integration of 20,000 unique sea urchin EST clusters
Genome Res. 13 (12), 2736-2746 (2003)
Contact: Poustka AJ

laboraty 145, dept.Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inhnstr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: poustka@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting
(ONF) to reduce sequencing redundancy. According to the ONF
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONF cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of
the other clones assigned to the same ONF cluster as the clone from
which the above EST is generated is available at the sea urchin
project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA
clones and filters are distributed via the Resource Center/Primary
Database of the German Human Genome Project (<http://www.rzpd.de>)
PCR Primers
FORWARD: 5' CCCAGCGTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GTATTACCGCAGCTGGCGAAAGGGGATGTG 3' (M13FSP) 3'-seq
Seq primer: 5'-CCGTCGGAATTCCTCCGGT-3' pSPORT3/86
High quality sequence stop: 399.

FEATURES
source
1..399
Location/Qualifiers
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="MPMGp691G19122:MPI_SURUDI_122G19"
/tissue_type="whole larva"
/dev_stage="larva 2-3 weeks"
/lab_host="E.coli, XL1 blue"
/clone_lib="Sea urchin larva cDNA library MPMGP691"
/notes="Vector: pSPORT1, Site 1: NotI; Site 2: SalI; Random
primed and directionally cloned in pSPORT1 vector using a
NotI (5'-PGACTAGTTTCTAGATCGGAGCGCGCC (T)15-3' and a
SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."

ORIGIN
Query Match 70.0%; Score 21; DB 6; Length 399;
Best Local Similarity 73.3%; Pred. No. 4e+02;
Mismatches 22; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GACAUGAACAAAGAGAGAUUAGGACAGAGT 30

Db 99 GACANGAGAGAGAGAGATTAGGAAGAAGT 128
|||||:|||||:|:|||||:|||||
|||||:|||||:|:|||||:|||||

RESULT 10
BF103720
LOCUS
DEFINITION
601647304F1 NTH_MGC_61 Homo sapiens cDNA clone IMAGE:3931440 5',
mRNA sequence.
BF103720
BF103720.1 GI:10886246

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 519)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCN767 row: h column: 01
High quality sequence stop: 518.

FEATURES
source
1..519
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3931440"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 61"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgcc); Site_2: SfiI (ggccatagcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCATTTATGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGCGCATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 70.0%; Score 21; DB 2; Length 519;
Best Local Similarity 75.9%; Pred. No. 4.2e+02;
Mismatches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 GACAUACAACAGAGAGAUUAGGACAGAGG 29
|||||:|||||:|:|||||:|||||
Db 488 GAAAGAACAGAGAGAGATTAAAGCAGAGG 516

RESULT 11
LOCUS
DEFINITION
Forward strand read from insert in 3'HPRT genomic targeting and
chromosome engineering clone MHP323e21, genomic survey sequence.
CR048521
CR048521.1 GI:49781660
GSS; genome survey sequence; MICE.
Mus musculus (house mouse)
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 693)
 Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
 Direct Submission
 CB10 1SA, UK. <http://www.sanger.ac.uk/MICR>
 Location/Qualifiers
 1..693
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHPP323e21"
 /clone_lib="MHPP"

TITLE
 JOURNAL
 FEATURES
 source

ORIGIN
 Query Match 70.0%; Score 21; DB 9; Length 693;
 Best Local Similarity 72.4%; Pred. No. 4.4e+02;
 Matches 21; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ACAGAACACAGAGAGAUUAGGCAGG 30
 Db 454 ACATGAAAAGAGATGCTAGGAAAGGT 426

RESULT 12
 BG104640/c
 LOCUS
 DEFINITION 602311331F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4421191 5', mRNA 880 bp EST 30-JAN-2001
 mRNA sequence.
 BG104640
 VERSION
 BG104640.1 GI:12598486
 EST.
 SOURCE
 Homo sapiens (human)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 880)
 NIH-MGC <http://mgs.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-k@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10159 row: n column: 08
 High quality sequence stop: 680.
 Location/Qualifiers
 1..880
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4421191"
 /tissue_type="adrenal cortex carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_84"
 Note: Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 70.0%; Score 21; DB 4; Length 880;
 Best Local Similarity 79.3%; Pred. No. 4.5e+02;
 Matches 23; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUUAGGCAGG 29
 Db 98 GAAAGGAACAAGAGGTGAGAAGGCAGG 70

RESULT 13
 CC258547
 LOCUS
 DEFINITION CH261-63K24 RM1.1 CH261 Gallus gallus genomic clone CH261-63K24, genomic survey sequence.
 CC258547
 VERSION
 CC258547.1 GI:30599491
 GSS.
 SOURCE
 Gallus gallus (chicken)
 ORGANISM
 Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 1 (bases 1 to 1110)
 Krenitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
 Gallus gallus BAC End Reads
 Unpublished (2003)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Insert Length: 182000 Std Error: 0.00
 Seq primer: RM1 TACGACTCACTATAGGAGAG
 Class: BAC ends
 High quality sequence start: 43
 High quality sequence stop: 696.
 Location/Qualifiers
 1..1110
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /strains="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-63K24"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /clone_lib="CH261"
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI; CH261 Female Chicken library - For library and clone ordering information: <http://www.chori.org/bacpac>"

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source

ORIGIN

Query Match 70.0%; Score 21; DB 8; Length 1110;
 Best Local Similarity 72.4%; Pred. No. 4.7e+02;
 Matches 21; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUUAGGCAGG 29
 Db 407 GAAATGAAGAAGATAAATAGGCAGG 435

RESULT 14
 BF342749
 LOCUS
 DEFINITION 602015013F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4150755 5', mRNA sequence.
 BF342749
 VERSION
 BF342749.1 GI:11289773
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1154)
 NIH-MGC <http://mgs.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

ORIGIN
 Query Match 70.0%; Score 21; DB 4; Length 880;
 Best Local Similarity 79.3%; Pred. No. 4.5e+02;
 Matches 23; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9414 row: j column: 04
High quality sequence stop: 657.

FEATURES
source

1. .1154
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4150755"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP_Brn64"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 70.0%; Score 21; DB 2; Length 1154;
Best Local Similarity 69.0%; Pred. No. 4.7e+02;
Matches 20; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGUAGUAGGCAGG 29

|||||:|||||:|:|||||

DB 328 GACCTGCAGCAGAGATTATTGGCAGG 356

RESULT 15

CG623002

LOCUS

CG623002 73 bp mRNA linear GSS 02-OCT-2003
OST323776 Mus musculus 129Sv/Ev Mus musculus cDNA clone OST323776,
mRNA sequence.

CG623002

CG623002.1 GI:37446851

GSS.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 73)

Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,

Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,

Fridde,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,

Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,

Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,

Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,

Zhu,Q., Person,C. and Sands,A.T.

Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap

screen to identify potential targets for therapeutic intervention

Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

Contact: Zambrowicz BP

Omnibank

Lexicon Genetics Incorporated

4000 Research Forest Drive, The Woodlands, TX 77381, USA

Email: materials@lexgen.com

Gene trap sequence tag generated by 3' RACE from mouse ES cells as

described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)

Class: Gene Trap.

Location/Qualifiers

1. .73

/organism="Mus musculus"

/mol_type="mRNA"

/strain="129Sv/Ev"

/db_xref="taxon:10090"

/clone="OST323776"

FEATURES
source

/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129Sv/Ev"

ORIGIN

Query Match 69.3%; Score 20.8; DB 9; Length 73;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 18; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 ACAUGAACAGAGAGUAGUAGGCAGG 28

|||||:|||||:|:|||||

DB 19 ACATGNACCAGAGATNATTTGGCAGG 45

Search completed: March 17, 2005, 11:07:42
Job time : 2082.4 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-60
Perfect score: 20
Sequence: 1 gacatgaacaagagatgatt 20
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004as: *
13: Geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	2 AAT72566	Aat72566 Hepatitis
2	20	100.0	20	2 AAT72565	Aat72565 Hepatitis
3	20	100.0	30	2 AAT72562	Aat72562 Hepatitis
4	20	100.0	30	2 AAT72563	Aat72563 Hepatitis
5	20	100.0	30	2 AAT72616	Aat72616 Hepatitis
6	20	100.0	30	2 AAT72617	Aat72617 Hepatitis
C 7	20	100.0	39	10 ADC64742	ADC64742 Hepatitis
C 8	20	100.0	87	4 AAD09094	Aad09094 Hepatitis
C 9	20	100.0	129	4 AAD09093	Aad09093 Hepatitis
C 10	20	100.0	639	6 AAD27422	Aad27422 Hepatitis
C 11	20	100.0	639	6 AAD31509	Aad31509 Hepatitis
C 12	20	100.0	655	4 AAT77569	Aah77569 HBV genot
C 13	20	100.0	655	4 AAT77568	Aah77568 HBV genot
C 14	20	100.0	655	4 AAT77574	Aah77574 HBV genot
C 15	20	100.0	655	4 AAT77573	Aah77573 HBV genot
C 16	20	100.0	655	4 AAT77570	Aah77570 HBV genot
C 17	20	100.0	655	4 AAT77571	Aah77571 HBV genot
C 18	20	100.0	664	4 AAT77572	Aah77572 HBV genot
C 19	20	100.0	669	12 AAD007220	Aad007220 Hepatitis
C 20	20	100.0	673	4 AAD09092	Aad09092 Hepatitis

C 21	20	100.0	675	4 AAT77563	Aah77563 HBV preCo
C 22	20	100.0	681	4 AAT77567	Aah77567 HBV genot
C 23	20	100.0	1395	2 AAV82688	Aav82688 Fulminant
C 24	20	100.0	1400	2 AAV82687	Aav82687 Fulminant
C 25	20	100.0	1445	2 AAV82692	Aav82692 Fulminant
C 26	20	100.0	1445	2 AAV82685	Aav82685 Fulminant
C 27	20	100.0	1445	2 AAV82690	Aav82690 Fulminant
C 28	20	100.0	1445	2 AAV82684	Aav82684 Fulminant
C 29	20	100.0	1500	2 AAV82695	Aav82695 Fulminant
C 30	20	100.0	1500	2 AAV82683	Aav82683 Fulminant
C 31	20	100.0	1500	2 AAV82694	Aav82694 Fulminant
C 32	20	100.0	1500	2 AAV82686	Aav82686 Fulminant
C 33	20	100.0	1500	2 AAV82706	Aav82706 Wild type
C 34	20	100.0	1500	2 AAV82689	Aav82689 Fulminant
C 35	20	100.0	1500	2 AAV82693	Aav82693 Fulminant
C 36	20	100.0	2342	1 AAN93072	Aan93072 Sequence
C 37	20	100.0	2743	1 AAN00003	Aan00003 Sequence
C 38	20	100.0	2743	2 AAQ04799	AAQ04799 Recombina
C 39	20	100.0	3180	4 AAH42375	AAH42375 Nucleotid
C 40	20	100.0	3182	6 AAD31765	AAH31765 Hepatitis
C 41	20	100.0	3182	9 ACA62422	ACA62422 Hepatitis
C 42	20	100.0	3182	10 AAD60866	AAD60866 Hepatitis
C 43	20	100.0	3220	3 AA288924	AA288924 Hepatitis
C 44	20	100.0	3248	4 AAD09091	Aad09091 Hepatitis
C 45	20	100.0	3248	4 AAT77562	Aah77562 HBV genot

ALIGNMENTS

RESULT 1									
AAT72566									
ID	AAT72566	standard; RNA; 20 BP.							
XX									
AC	AAT72566;								
DT	03-SEP-1997	(first entry)							
XX									
DE	Hepatitis B virus RNA antisense oligonucleotide HBV46MYb.								
KW	HBV; HBV infection; inhibition; replication; ss.								
XX									
OS	Synthetic.								
XX									
FH	Key	Location/Qualifiers							
FT	misc_feature	1..20							
FT		/tag= a							
FT		/note= "Internucleotide linkages are phosphorothioate"							
FT	modified_base	1							
FT		/tag= b							
FT	modified_base	2							
FT		/mod_base= gm							
FT		/tag= c							
FT		/mod_base= OTHER							
FT	modified_base	3							
FT		/note= "2'-O-methyladenosine"							
FT		/tag= d							
FT	modified_base	4							
FT		/mod_base= cm							
FT		/tag= e							
FT		/mod_base= OTHER							
FT	modified_base	5							
FT		/note= "2'-O-methyladenosine"							
FT		/tag= f							
FT	modified_base	6							
FT		/mod_base= um							
FT		/tag= g							
FT	modified_base	7							
FT		/mod_base= gm							
FT		/tag= h							
FT		/mod_base= OTHER							
FT	modified_base	8							
FT		/note= "2'-O-methyladenosine"							

```

FT FT /*tag= i
FT FT /mod_base= OTHER
FT FT 9 /note= "2'-O-methyladenosine"
FT FT modified_base
FT FT /*tag= j
FT FT 10 /mod_base= cm
FT FT modified_base
FT FT /*tag= k
FT FT 11 /mod_base= OTHER
FT FT 12 /note= "2'-O-methyladenosine"
FT FT modified_base
FT FT /*tag= l
FT FT 13 /mod_base= OTHER
FT FT 14 /note= "2'-O-methyladenosine"
FT FT modified_base
FT FT /*tag= m
FT FT 15 /mod_base= gm
FT FT modified_base
FT FT /*tag= n
FT FT 16 /mod_base= OTHER
FT FT 17 /note= "2'-O-methyladenosine"
FT FT modified_base
FT FT /*tag= o
FT FT 18 /mod_base= gm
FT FT modified_base
FT FT /*tag= p
FT FT 19 /mod_base= OTHER
FT FT 20 /note= "2'-O-methyladenosine"
FT FT modified_base
FT FT /*tag= q
FT FT 21 /mod_base= um
FT FT modified_base
FT FT /*tag= r
FT FT 22 /mod_base= gm
FT FT modified_base
FT FT /*tag= s
FT FT 23 /mod_base= OTHER
FT FT 24 /note= "2'-O-methyladenosine"
FT FT modified_base
FT FT /*tag= t
FT FT 25 /mod_base= um
FT FT modified_base
FT FT /*tag= u
FT FT 26 /mod_base= um
FT FT modified_base
FT FT WO9639502-A1.
FT FT 12-DEC-1996.
FT FT 04-JUN-1996; 96WO-EP002432.
FT FT 06-JUN-1995; 95US-00467397.
FT FT (HOFF ) HOFFMANN LA ROCHE & CO AG F.
FT FT (HYBR-) HYBRIDON INC.
FT FT Craig CU, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
FT FT Roberts NA, Roberts PC, Slade A;
FT FT WPI; 1997-043124/04.
FT FT
FT FT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
FT FT used in the detection and treatment of HBV infection.
FT FT
FT FT Claim 1; Page 12; 81pp; English.
FT FT
FT FT The present sequence represents a synthetic oligonucleotide HBV46MYB
FT FT which is complementary to a portion of the hepatitis B virus (HBV) RNA.
FT FT The antisense oligonucleotide may be used to detect the presence of HBV
FT FT in a sample. The antisense oligonucleotide, and oligonucleotides
FT FT containing a sequence which is complementary to at least two non-
FT FT contiguous regions of an HBV nucleic acid, may be used for inhibiting HBV
FT FT replication in a cell or for the treatment of HBV infection

```

```

XX SQ Sequence 20 BP; 9 A; 2 C; 5 G; 0 T; 4 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 9.6;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
Db 1 GACAUGAACAAGAGAGAUU 20

RESULT 2
AAT72565
ID AAT72565 standard; DNA; 20 BP.
XX AC AAT72565;
DT 03-SEP-1997 (first entry)
DE Hepatitis B virus RNA antisense oligonucleotide HBV46Yb.
XX HBV; HBV infection; inhibition; replication; ss.
XX Synthetic.
XX Key Location/Qualifiers
FT misc_feature 1..20
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
XX WO9639502-A1.
XX 12-DEC-1996.
XX 04-JUN-1996; 96WO-EP002432.
XX 06-JUN-1995; 95US-00467397.
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX (HYBR-) HYBRIDON INC.
XX Craig CU, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
XX Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
XX used in the detection and treatment of HBV infection.
XX
XX Claim 1; Page 12; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV46Yb which
XX is complementary to a portion of the hepatitis B virus (HBV) RNA. The
XX antisense oligonucleotide may be used to detect the presence of HBV in a
XX sample. The antisense oligonucleotide, and oligonucleotides containing a
XX sequence which is complementary to at least two non- contiguous regions
XX of an HBV nucleic acid, may be used for inhibiting HBV replication in a
XX cell or for the treatment of HBV infection
XX SQ Sequence 20 BP; 9 A; 2 C; 5 G; 4 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
Db 1 GACATGAACAAGAGATGATT 20

RESULT 3
AAT72562

```

ID AAT72562 standard; DNA; 30 BP.
 AC AAT72562;
 XX
 DT 03-SEP-1997 (first entry)
 DE Hepatitis B virus RNA antisense oligonucleotide HBV88b.
 XX
 KW HBV; HBV infection; inhibition; replication; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 1..30
 FT /tag= a
 FT /note= "Internucleotide linkages are phosphorothioate"
 XX
 PN WO9639502-A1.
 PD 12-DEC-1996.
 XX
 PP 04-JUN-1996; 96WO-EP002432.
 XX
 PR 06-JUN-1995; 95US-00467397.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA (HYBR-) HYBRIDON INC.
 XX
 PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
 PI Roberts NA, Roberts PC, Slade A;
 XX
 DR WPI; 1997-043124/04.
 XX
 PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 PT used in the detection and treatment of HBV infection.
 XX
 PS Claim 1; Page 12; 81pp; English.
 XX
 CC The present sequence represents a synthetic oligonucleotide HBV88b which
 CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
 CC antisense oligonucleotide may be used to detect the presence of HBV in a
 CC sample. The antisense oligonucleotide, and oligonucleotides containing a
 CC sequence which is complementary to at least two non- contiguous regions
 CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
 CC cell or for the treatment of HBV infection
 XX
 SQ Sequence 30 BP; 12 A; 3 C; 10 G; 5 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 30;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGACACAGAGATGATT 20
 DB 1 GACATGACACAGAGATGATT 20
 RESULT 4
 AAT72563
 ID AAT72563 standard; DNA; 30 BP.
 XX
 AC AAT72563;
 XX
 DT 03-SEP-1997 (first entry)
 DE Hepatitis B virus RNA antisense oligonucleotide HBV88Mb.
 XX
 KW HBV; HBV infection; inhibition; replication; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 1..30

FT /tag= a
 FT /note= "Internucleotide linkages are phosphorothioate"
 FT 1..20
 FT /tag= b
 FT /note= "2'-O-Me RNA"
 FT modified_base
 FT /tag= c
 FT /mod_base= gm
 FT modified_base
 FT /tag= d
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= e
 FT /mod_base= cm
 FT modified_base
 FT /tag= f
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= g
 FT /mod_base= um
 FT modified_base
 FT /tag= h
 FT /mod_base= gm
 FT modified_base
 FT /tag= i
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= j
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= k
 FT /mod_base= cm
 FT modified_base
 FT /tag= l
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= m
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= n
 FT /mod_base= gm
 FT modified_base
 FT /tag= o
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= p
 FT /mod_base= gm
 FT modified_base
 FT /tag= q
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= r
 FT /mod_base= um
 FT modified_base
 FT /tag= s
 FT /mod_base= gm
 FT modified_base
 FT /tag= t
 FT /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 FT modified_base
 FT /tag= u
 FT /mod_base= um
 FT modified_base
 FT /tag= v

```

FT XX                               /mod_base= um
PN XX
XX WO9639502-A1.
XX
PD 12-DEC-1996.
XX
XX 04-JUN-1996; 96WO-EP002432.
PF
XX
XX 06-JUN-1995; 95US-00467397.
PR
XX
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
PA
XX (HYBR-) HYBRIDON INC.
XX
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
FT used in the detection and treatment of HBV infection.
XX
XX Claim 1; Page 12; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV89Mb which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a
CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non- contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection
XX
XX Sequence 30 BP; 12 A; 3 C; 10 G; 1 T; 4 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 80.0%; Pred. No. 9.9;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
Db |||:|||||:|||||:|:|:|
1 GACAUGAACAAAGAGAGUAU 20

RESULT 5
AAT72616
ID AAT72616 standard; DNA; 30 BP.
XX
XX AC AAT72616;
XX
XX 04-SEP-1997 (first entry)
DT
DE Hepatitis B virus RNA antisense oligonucleotide HBV-89b.
XX
XX HBV; HBV infection; inhibition; replication; ss.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH misc_feature 1..30
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
FT misc_RNA 1..20
FT /*tag= b
FT /note= "2'-Ome RNA"
FT modified_base 1
FT /*tag= c
FT /mod_base= gm
FT modified_base 2
FT /*tag= d
FT /mod_base= OTHER
FT modified_base 3
FT /note= "2'-O-methyladenosine"
FT /*tag= e
FT modified_base 4
FT /mod_base= cm
FT /*tag= f
FT /mod_base= OTHER
FT modified_base 5
FT /note= "2'-O-methyladenosine"
FT /*tag= g
FT /mod_base= um
FT modified_base 6
FT /*tag= h
FT /mod_base= gm
FT modified_base 7
FT /*tag= i
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"

```

```

PI Roberts NA, Roberts PC, Slade A;
XX
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
FT used in the detection and treatment of HBV infection.
XX
XX Claim 5; Page 15; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV-89b which
CC contains a sequence which is complementary to at least two non-contiguous
CC regions of a hepatitis B virus (HBV) nucleic acid. The antisense
CC oligonucleotide may be used to detect the presence of HBV in a sample.
CC The antisense oligonucleotide, and oligonucleotides complementary to a
CC portion of the HBV RNA, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection
XX
XX Sequence 30 BP; 12 A; 3 C; 9 G; 6 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
Db |||:|||||:|||||:|:|:|
1 GACATGAACAAGAGATGATT 20

RESULT 6
AAT72617
ID AAT72617 standard; DNA; 30 BP.
XX
XX AC AAT72617;
XX
XX 04-SEP-1997 (first entry)
DT
DE Hepatitis B virus RNA antisense oligonucleotide HBV-89Mb.
XX
XX HBV; HBV infection; inhibition; replication; ss.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH misc_feature 1..30
FT /*tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
FT misc_RNA 1..20
FT /*tag= b
FT /note= "2'-Ome RNA"
FT modified_base 1
FT /*tag= c
FT /mod_base= gm
FT modified_base 2
FT /*tag= d
FT /mod_base= OTHER
FT modified_base 3
FT /note= "2'-O-methyladenosine"
FT /*tag= e
FT modified_base 4
FT /mod_base= cm
FT /*tag= f
FT /mod_base= OTHER
FT modified_base 5
FT /note= "2'-O-methyladenosine"
FT /*tag= g
FT /mod_base= um
FT modified_base 6
FT /*tag= h
FT /mod_base= gm
FT modified_base 7
FT /*tag= i
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"

```

```
FT modified_base 8 /tag= j
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 9 /tag= k
FT /mod_base= cm
FT modified_base 10 /tag= l
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 11 /tag= m
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 12 /tag= n
FT /mod_base= gm
FT modified_base 13 /tag= o
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 14 /tag= p
FT /mod_base= gm
FT modified_base 15 /tag= q
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 16 /tag= r
FT /mod_base= um
FT modified_base 17 /tag= s
FT /mod_base= gm
FT modified_base 18 /tag= t
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 19 /tag= u
FT /mod_base= um
FT modified_base 20 /tag= v
FT /mod_base= um
FT
FT WO9639502-A1.
FT
FT 12-DEC-1996.
FT
FT 04-JUN-1996; 96WO-EP002432.
FT
FT 06-JUN-1995; 95US-00467397.
FT
FT (HOFF ) HOFFMANN LA ROCHE & CO AG F.
FT (HYBR-) HYBRIDON INC.
FT
FT Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
FT Roberts NA, Roberts PC, Slade A;
FT WPI; 1997-043124/04.
FT
FT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
FT used in the detection and treatment of HBV infection.
FT
FT Claim 5; Page 15; 81pp; English.
FT
FT The present sequence represents a synthetic oligonucleotide HBV-89Mb
FT which contains a sequence which is complementary to at least two non-
FT contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
FT antisense oligonucleotide may be used to detect the presence of HBV in a
FT sample. The antisense oligonucleotide, and oligonucleotides complementary
FT to a portion of the HBV RNA, may be used for inhibiting HBV replication
```

```
CC in a cell or for the treatment of HBV infection
XX
SQ Sequence 30 BP; 12 A; 3 C; 9 G; 2 T; 4 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 80.0%; Pred. No. 9.9;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATGAACAAGAGATGATT 20
Db 1 GACAUGAACACAGAGAUU 20
RESULT 7
ADC64742/c
ID ADC64742 standard; RNA; 39 BP.
XX
AC ADC64742;
XX
DT 18-DEC-2003 (first entry)
XX
DE Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX
KW screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX
OS Synthetic.
OS Hepatitis B virus.
XX
PN KR2002007891-A.
XX
PD 29-JAN-2002.
XX
PF 19-JUL-2000; 2000KR-00041420.
XX
PR 19-JUL-2000; 2000KR-00041420.
XX
PA (MOGA-) MOGAM BIOTECHNOLOGY INST.
PA (VIRO-) VIROGEN CO LTD.
XX
PI Ji HJ, Jung SI, Kim YC, Min MG, Ryu WS, Yoon GS;
XX
DR WPI; 2003-309015/30.
XX
PT Screening of antiviral agents by protein-priming activity of hepatitis B
PT virus DNA polymerase.
XX
PS Disclosure; Page 12; 13pp; Korean.
XX
CC The present invention describes a method of screening for an antiviral
CC agent by the protein-priming activity of hepatitis B virus (HBV) DNA
CC polymerase. Also described is developing an antiviral agent with a high
CC selectivity to HBV which can be used for high-throughput screening. The
CC present sequence represents an RNA oligonucleotide which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 39 BP; 5 A; 13 C; 3 G; 0 T; 18 U; 0 Other;
Query Match 100.0%; Score 20; DB 10; Length 39;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATGAACAAGAGATGATT 20
Db 37 GACATGAACAAGAGATGATT 18
RESULT 8
AAD09094/c
ID AAD09094 standard; DNA; 87 BP.
XX
AC AAD09094;
XX
DT 04-SEP-2001 (first entry)
```

```

XX DE Hepatitis B virus FRI strain genotype G HBeAg DNA fragment.
XX OS
XX KW HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
XX PN preS2; surface antigen; HBeAg; HBx protein; vaccine; HBeAg;
XX KW liver disease; hepatitis; liver cancer; HBeAg; core antigen; ds.
XX OS
XX OS Hepatitis B virus.
XX PN WO200138498-A2.
XX XX 31-MAY-2001.
XX PF 21-NOV-2000; 2000WO-US032108.
XX PN 24-NOV-1999; 99US-0167206P.
XX XX (PHAR-) PHARMASSET INC.
XX PA (INNO-) INNOGENETICS NV.
XX PI Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M;
XX PI Rossau R;
XX DR WPI; 2001-367676/38.
XX XX
XX PT Novel hepatitis B virus genotype G, nucleic acids encoding virus,
XX PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
XX PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX PS Claim 6; Page 57; 84pp; English.
XX CC The present invention relates to hepatitis B virus (HBV) strain FRI,
XX CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
XX CC PreS2 and surface antigen HBeAg) and HBx proteins. HBV genotype G nucleic
XX CC acids and polypeptides are useful for diagnosing, prognosing and treating
XX CC infections caused by HBV genotype G. They can be used in a vaccine to
XX CC treat or prevent HBV genotype G infection. The HBV genotype G derived
XX CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
XX CC sample or diagnosis of HBV genotype G infection. The presence of HBV
XX CC genotype G statistically correlates with the presence of liver damage
XX CC and/or liver cancer in the subject. The HBV genotype G core insert
XX CC peptide encoding nucleic acid is useful for designing monitoring assays
XX CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
XX CC and HBeAg (genotype G e antigen) in patients infected with HBV. The
XX CC antibodies or antigens of HBV genotype G are useful for identifying a
XX CC stage of liver disease caused by HBV genotype G. The present sequence is
XX CC a hepatitis B virus (HBV) strain FRI, genotype G DNA fragment encoding e
XX CC antigen (HBeAg)
XX SX
XX SQ Sequence 87 BP; 14 A; 24 C; 17 G; 32 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 87;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
DB 43 GACATGAACAAGAGATGATT 24

RESULT 9
AAD09093/C
ID AAD09093 standard; DNA; 129 BP.
XX AC AAD09093;
XX DT 04-SEP-2001 (first entry)
XX XX
DE Hepatitis B virus FRI strain genotype G DNA fragment #1.
XX KW HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
XX KW preS2; surface antigen; HBeAg; HBx protein; vaccine; liver disease;
XX KW hepatitis; liver cancer; HBeAg; core antigen; ds.

```

```

XX OS Hepatitis B virus.
XX PN WO200138498-A2.
XX XX 31-MAY-2001.
XX PF 21-NOV-2000; 2000WO-US032108.
XX PN 24-NOV-1999; 99US-0167206P.
XX XX (PHAR-) PHARMASSET INC.
XX PA (INNO-) INNOGENETICS NV.
XX PI Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M;
XX PI Rossau R;
XX DR WPI; 2001-367676/38.
XX XX
XX PT Novel hepatitis B virus genotype G, nucleic acids encoding virus,
XX PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
XX PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX PS Claim 5; Page 57; 84pp; English.
XX CC The present invention relates to hepatitis B virus (HBV) strain FRI,
XX CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
XX CC PreS2 and surface antigen HBeAg) and HBx proteins. HBV genotype G nucleic
XX CC acids and polypeptides are useful for diagnosing, prognosing and treating
XX CC infections caused by HBV genotype G. They can be used in a vaccine to
XX CC treat or prevent HBV genotype G infection. The HBV genotype G derived
XX CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
XX CC sample or diagnosis of HBV genotype G infection. The presence of HBV
XX CC genotype G statistically correlates with the presence of liver damage
XX CC and/or liver cancer in the subject. The HBV genotype G core insert
XX CC peptide encoding nucleic acid is useful for designing monitoring assays
XX CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
XX CC and HBeAg (genotype G e antigen) in patients infected with HBV. The
XX CC antibodies or antigens of HBV genotype G are useful for identifying a
XX CC stage of liver disease caused by HBV genotype G. The present sequence is
XX CC a hepatitis B virus (HBV) strain FRI, genotype G DNA fragment
XX SQ Sequence 129 BP; 25 A; 32 C; 26 G; 46 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
DB 43 GACATGAACAAGAGATGATT 24

RESULT 10
AAD27422/C
ID AAD27422 standard; DNA; 639 BP.
XX AC AAD27422;
XX DT 18-APR-2002 (first entry)
XX XX
DE Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
XX KW Hepatitis B virus; HBV; core antigen; HBeAg; immune system; typhoid;
XX KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
XX KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
XX KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
XX KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
XX KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
XX KW antiprotocool; ds.
XX OS Hepatitis B virus.

```

```

FH Key      Location/Qualifiers
FT CDS      1..639
FT FT       /*tag= a
FT FT       /product= "HBc protein"
FT FT       1..87
FT FT       /*tag= b
FT FT       88..636
FT FT       /*tag= c
FT FT       /product= "Mature HBc protein"
PN WO200198333-A2.
XX 27-DEC-2001.
XX
XX 22-JUN-2001; 2001WO-GB002817.
XX
XX 22-JUN-2000; 2000GB-00015308.
XX
XX 06-OCT-2000; 2000GB-00024544.
XX
XX (CELL-) CELTECH PHARM LTD.
XX
XX Page M, Li J, Pumpens P;
XX
XX WPI; 2002-098223/13.
XX
XX P-PSDB; AAE17018.
XX
XX New proteins comprising a modified hepatitis B core antigen, useful as a
XX vaccine in prophylactic or therapeutic vaccination of the human or animal
XX body, particularly against hepatitis B virus infection.
XX
XX Disclosure; Page 38-39; 40pp; English.
XX
XX The invention relates to modified proteins comprising hepatitis B virus
XX (HBV) core antigen (HBcAg) wherein one or more of the four arginine
XX repeats has been deleted and the protein comprising the C-terminal
XX cysteine of HBcAg. The deleted region may be replaced by an epitope from
XX a protein other than HBcAg, in which case the HBcAg acts as a carrier to
XX present the epitope to the immune system. This chimeric protein or its
XX nucleic acid is useful as a vaccine or in a method of prophylactic or
XX therapeutic vaccination of the human or animal body, particularly against
XX HBV. The nucleic acid encoding the protein may be used in gene therapy or
XX DNA vaccination protocols. The chimeric protein or its nucleic acid may
XX also be used as the basis of a prophylactic vaccine against a range of
XX diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
XX influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
XX immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
XX tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
XX diarrhoea, meningitis or gonorrhoea. The present sequence is a DNA
XX encoding Hepatitis B virus core antigen (HBcAg)
XX
XX Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

Query Match      100.0%; Score 20; DB 6; Length 639;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
Db |||||
43 GACATGAACAAGAGATGATT 24

RESULT 11
AAD31509/c
ID AAD31509 standard; DNA; 639 BP.
XX
XX AAD31509;
XX
XX 18-JUN-2002 (first entry)
XX
XX Hepatitis B virus core antigen (HBcAg) encoding DNA.
XX
XX Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
XX therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
XX polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX
XX Hepatitis B virus.
XX
XX Key      Location/Qualifiers

FH CDS      1..639
FT FT       /*tag= a
FT FT       /product= "HBc protein"
FT FT       1..87
FT FT       /*tag= b
FT FT       88..636
FT FT       /*tag= c
FT FT       /product= "Mature HBc protein"
PN WO200177158-A1.
XX
XX 18-OCT-2001.
XX
XX 09-APR-2001; 2001WO-GB001607.
XX
XX 07-APR-2000; 2000EP-00107118.
XX
XX (MEDE-) MEDEVA EURO LTD.
XX
XX Gehin A, Gilbert R, Stuart D, Rowlands D;
XX
XX WPI; 2002-239995/29.
XX
XX P-PSDB; AAE19793.
XX
XX Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
XX prophylactic or therapeutic treatment of humans or animals against e.g.
XX HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
XX mouth disease.
XX
XX Disclosure; Page 23-24; 27pp; English.
XX
XX The present invention relates to hepatitis B virus (HBV) core antigen
XX (HBcAg) fusion proteins and polynucleotides encoding such proteins.
XX Sequences of the invention are useful in methods of prophylactic or
XX therapeutic vaccination or to manufacture medicaments for prophylactic or
XX therapeutic vaccination of the human or animal body against HBV, e.g.
XX against viral hepatitis. They are also useful as a prophylactic vaccine
XX against e.g. hepatitis c virus, influenza, polio, herpes, rabies,
XX acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
XX present sequence is a DNA encoding hepatitis B virus core antigen (HBcAg)
XX
XX Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

Query Match      100.0%; Score 20; DB 6; Length 639;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGATT 20
Db |||||
43 GACATGAACAAGAGATGATT 24

RESULT 12
AAH77569/c
ID AAH77569 standard; DNA; 655 BP.
XX
XX AAH77569;
XX
XX 19-OCT-2001 (first entry)
XX
XX HBV genotype G strain US1 preCore/Core DNA.
XX
XX Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HBSPol;
XX HBsAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBcAg;
XX HBsAg; ds.
XX
XX Hepatitis B virus.
XX
XX WO200140279-A2.
XX
XX 07-JUN-2001.
XX
XX 20-NOV-2000; 2000WO-BF011526.

```

XX 03-DEC-1999; 99EP-00870252.
 PR 07-DEC-1999; 99US-0169287P.
 XX (INNO-) INNOGENETICS NV.
 XX Stuyver L, Van Geyt C, De Gendt S;
 PI WPI; 2001-374785/39.
 XX Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.
 XX Claim 3; Fig 7; 94pp; English.
 PS The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
 CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for HBV genotyping. The proteins encoded by
 CC the polynucleotides are useful for detecting antibodies in a biological
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBeAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is provided in an alignment of preCore/Core sequences of
 CC an HBV genotype A strain (HBVXCPs) and 7 strains (FR1, FR2, US1, US3,
 CC US6, US7, US9, US10) of HBV genotype G
 XX Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;
 SQ Query Match 100.0%; Score 20; DB 4; Length 655;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGACACAGAGATGATT 20
 DB |||||
 43 GACATGACACAGAGATGATT 24
 RESULT 13
 AAH77568/c
 ID AAH77568 standard; DNA; 655 BP.
 XX AAH77568;
 AC 19-OCT-2001 (first entry)
 XX HBV genotype G strain FR2 preCore/Core DNA.
 DE Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBx; HBPol;
 XX HBSAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBeAg;
 KW HBeAg; ds.
 XX Hepatitis B virus.
 OS WO200140279-A2.
 XX 07-JUN-2001.
 PD 20-NOV-2000; 2000WO-EP011526.
 XX 03-DEC-1999; 99EP-00870252.
 PR 07-DEC-1999; 99US-0169287P.
 XX (INNO-) INNOGENETICS NV.
 PA Stuyver L, Van Geyt C, De Gendt S;
 PI WPI; 2001-374785/39.
 XX Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.

XX WPI; 2001-374785/39.
 DR Novel isolated and/or purified hepatitis B virus polypeptide and
 XX polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.
 XX Claim 3; Fig 7; 94pp; English.
 PS The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
 CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for HBV genotyping. The proteins encoded by
 CC the polynucleotides are useful for detecting antibodies in a biological
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBeAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is provided in an alignment of preCore/Core sequences of
 CC an HBV genotype A strain (HBVXCPs) and 7 strains (FR1, FR2, US1, US3,
 CC US6, US7, US9, US10) of HBV genotype G
 XX Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;
 SQ Query Match 100.0%; Score 20; DB 4; Length 655;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGACACAGAGATGATT 20
 DB |||||
 43 GACATGACACAGAGATGATT 24
 RESULT 14
 AAH77574/c
 ID AAH77574 standard; DNA; 655 BP.
 XX AAH77574;
 AC 19-OCT-2001 (first entry)
 XX HBV genotype G strain US10 preCore/Core DNA.
 DE Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBx; HBPol;
 XX HBSAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBeAg;
 KW HBeAg; ds.
 XX Hepatitis B virus.
 OS WO200140279-A2.
 XX 07-JUN-2001.
 PD 20-NOV-2000; 2000WO-EP011526.
 XX 03-DEC-1999; 99EP-00870252.
 PR 07-DEC-1999; 99US-0169287P.
 XX (INNO-) INNOGENETICS NV.
 PA Stuyver L, Van Geyt C, De Gendt S;
 PI WPI; 2001-374785/39.
 XX Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.

XX Claim 3; Fig 7; 94pp; English.
 XX The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
 CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for detecting the proteins and for detecting
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBeAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is provided in an alignment of preCore/Core sequences of
 CC an HBV genotype A strain (HBVXCPs) and 7 strains (FR1, FR2, US1, US3,
 CC US6, US7, US9, US10) of HBV genotype G
 XX
 SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;
 Query Match 100.0%; Score 20; DB 4; Length 655;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGAACAAGAGATGATT 20
 DB 43 GACATGAACAAGAGATGATT 24
 RESULT 15
 AAH77573/C
 ID AAH77573 standard; DNA; 655 BP.
 AC AAH77573;
 XX
 DT 19-OCT-2001 (first entry)
 DE HBV genotype G strain US7 preCore/Core DNA.
 KW Hepatitis B virus; HBV; preCore; Core; preS1; preS2; HBS; HBX; HSPol;
 KW HBeAg; antiviral; vaccine; genotype G; genotype A; genotyping; HBeAg;
 KW HBeAg; ds.
 OS Hepatitis B virus.
 XX
 PN WO200140279-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 20-NOV-2000; 2000WO-EP011526.
 XX
 PR 03-DEC-1999; 99EP-00870252.
 PR 07-DEC-1999; 99US-0169287P.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Stuyver L, Van Geyt C, De Gendt S;
 XX
 DR WPI; 2001-374785/39.
 XX
 PT Novel isolated and/or purified hepatitis B virus polypeptide and
 PT polynucleotide sequences that are phylogenetically different from HBV
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and
 PT therapy.
 XX
 PS Claim 3; Fig 7; 94pp; English.
 XX
 CC The invention relates to the complete nucleic acid sequence of a new
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype G.
 CC This genotype was found with a high prevalence in patients chronically
 CC infected with HBV and residing in Europe and the USA. The invention
 CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for detecting the proteins and for detecting
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBeAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is provided in an alignment of preCore/Core sequences of
 CC an HBV genotype A strain (HBVXCPs) and 7 strains (FR1, FR2, US1, US3,
 CC US6, US7, US9, US10) of HBV genotype G

CC relates to a fully defined sequence of 3248 nucleotides as given in
 CC specification, a sequence with 92% identity to the given sequence, or
 CC sequence that is degenerate to the mentioned sequences. These
 CC polynucleotides are useful for HBV genotyping. The proteins encoded by
 CC the polynucleotides are useful for detecting antibodies in a biological
 CC sample. Ligands that bind to the proteins and antibodies directed against
 CC the proteins are useful for detecting the proteins and for detecting
 CC HBeAg and HBeAg (precursor proteins). They are also useful for
 CC preparing a vaccine or medicament for treating HBV infections. The
 CC present sequence is provided in an alignment of preCore/Core sequences of
 CC an HBV genotype A strain (HBVXCPs) and 7 strains (FR1, FR2, US1, US3,
 CC US6, US7, US9, US10) of HBV genotype G
 XX
 SQ Sequence 655 BP; 144 A; 156 C; 143 G; 206 T; 0 U; 6 Other;
 Query Match 100.0%; Score 20; DB 4; Length 655;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACATGAACAAGAGATGATT 20
 DB 43 GACATGAACAAGAGATGATT 24
 Search completed: March 17, 2005, 06:48:41
 Job time : 171.333 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-60
Perfect score: 20
Sequence: 1 gacatgaacaagatgatt 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	18.4	92.0	763	9	CG381974	CG381974 OG1BK24TH
2	18.4	92.0	773	9	CG381984	CG381984 OG1BK24TH
3	18.4	92.0	846	9	CG373212	CG373212 OG1C228TH
4	18.4	92.0	892	9	CG373225	CG373225 OG1C228TV
5	17.4	87.0	280	9	BX288914	BX288914 Arabidops
6	17.4	87.0	305	6	CA520799	CA520799 KS1101781
7	17.4	87.0	313	7	CF906991	CF906991 A0504F09-
8	17.4	87.0	423	5	BX837806	BX837806 BX837806
9	17.4	87.0	426	8	AQ183243	AQ183243 HS 3140 B
10	17.4	87.0	442	7	CN958899	CN958899 6399.1001
11	17.4	87.0	450	8	AZ654527	AZ654527 IM0528D16
12	17.4	87.0	509	3	AK015232	AK015232 Mus muscu
13	17.4	87.0	527	8	AQ961278	AQ961278 LERFK20TF
14	17.4	87.0	570	7	CF198522	CF198522 EST0117 T
15	17.4	87.0	574	7	CF909462	CF909462 A0536H03-
16	17.4	87.0	616	7	CK517114	CK517114 rswj00 00
17	17.4	87.0	628	8	BH366000	BH366000 CH230-112
18	17.4	87.0	631	2	BE388774	BE388774 601283896
19	17.4	87.0	652	8	BZ898091	BZ898091 CH240.13N
20	17.4	87.0	682	9	CE163565	CE163565 tigr-gss-
21	17.4	87.0	683	7	CO817689	CO817689 FA SEA001
22	17.4	87.0	721	8	AQ961277	AQ961277 LERFK20TF
23	17.4	87.0	794	9	CL809904	CL809904 OR_Cha002
24	17.4	87.0	821	2	BF678287	BF678287 602084906

C	25	17.4	87.0	859	8	CC090167	CC090167 CSU-K33r.
C	26	17.4	87.0	870	8	CC131380	CC131380 NDL.76K8.
C	27	17.4	87.0	885	2	BF541940	BF541940 602068444
C	28	17.4	87.0	903	3	CC068329	CC068329 CSU-K33r.
C	29	17.4	87.0	1340	3	CNS0A5V5	EX823129 Arabidops
C	30	17	85.0	472	2	AW760013	AW760013 el56h09.y
C	31	17	85.0	532	1	AL819446	AL819446 AL819446
C	32	17	85.0	577	4	BI378081	BI378081 BFLG3_001
C	33	17	85.0	687	9	AG140746	AG140746 Pan trogl
C	34	17	85.0	730	5	EX114353	EX114353 BX114353
C	35	17	85.0	768	9	CC786410	CC786410 ZMMBB0015
C	36	17	85.0	778	7	CO368799	CO368799 RTK1.42.H
C	37	17	85.0	808	9	CL543431	CL543431 OB_Ba006
C	38	17	85.0	1013	9	CNS06RAQ	AL411720 T3 end of
C	39	16.8	84.0	124	4	BI127956	BI127956 G068P81Y
C	40	16.8	84.0	170	4	BI128183	BI128183 G072P24Y
C	41	16.8	84.0	195	4	BG125265	BG125265 EST470911
C	42	16.8	84.0	195	4	BG733602	BG733602 Cg-20d C1
C	43	16.8	84.0	218	2	BE428564	BE428564 MTD008.D0
C	44	16.8	84.0	234	4	BI473621	BI473621 fp39h03.y
C	45	16.8	84.0	251	8	BZ385056	BZ385056 SALK_1363

ALIGNMENTS

RESULT 1
CG381974 763 bp DNA linear GSS 26-AUG-2003
OG1BK24TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0724C23,
genomic survey sequence.
ACCESSION CG381974
VERSION CG381974.1 GI:34299241
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 763)
AUTHORS Whiteley, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG1BK24TV
COMMENT Contact: Cathy Whitelaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES

source
1..763
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0724C23"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 763;
Best Local Similarity 95.0%; Pred. No. 4.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GACATGAACAAGATGATT 20
DB 572 GGCATGAACAAGATGATT 591

```

RESULT 2
CG381984/c
LOCUS
DEFINITION CG381984 773 bp DNA linear GSS 26-AUG-2003
genomic survey sequence.
ACCESSION CG381984
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 773)
REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OG1BK24TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..773
Location/Qualifiers
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 92.0%; Score 18.4; DB 9; Length 773;
Best Local Similarity 95.0%; Pred. No. 4.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GACATGACACAGAGATGATT 20
DB 668 GGCATGACACAGAGATGATT 649
RESULT 3
CG373212/c
LOCUS
DEFINITION CG373212 846 bp DNA linear GSS 26-AUG-2003
genomic survey sequence.
ACCESSION CG373212
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 846)
REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OG1CZ28TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..892
Location/Qualifiers
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 92.0%; Score 18.4; DB 9; Length 892;
Best Local Similarity 95.0%; Pred. No. 4.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GACATGACACAGAGATGATT 20
DB 540 GGCATGACACAGAGATGATT 521
RESULT 4
CG373225
LOCUS
DEFINITION CG373225 892 bp DNA linear GSS 26-AUG-2003
genomic survey sequence.
ACCESSION CG373225
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 892)
REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OG1CZ28TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..892
Location/Qualifiers
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 92.0%; Score 18.4; DB 9; Length 892;
Best Local Similarity 95.0%; Pred. No. 4.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GACATGACACAGAGATGATT 20
DB 540 GGCATGACACAGAGATGATT 521

```

```

Db      353 GGCATGAACAAGAGATGATT 372

RESULT 5
BX288914/c
LOCUS      280 bp      DNA      linear      GSS 02-APR-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-421D09-018141,
             genomic survey sequence.
ACCESSION  BX288914
VERSION     BX288914.1 GI:28887910
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE   1
AUTHORS     Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weisshaar, B.
TITLE       GABI-Kat Simplesearch: a flanking sequence tag (fST) database for
             the identification of T-DNA insertion mutants in Arabidopsis
             thaliana
JOURNAL     Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE     22755829
PUBMED      12874060
REFERENCE   2
AUTHORS     Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and
             Weisshaar, B.
TITLE       An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
             flanking sequence tag-based reverse genetics
JOURNAL     Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE     23117147
PUBMED      14756321
REFERENCE   3
AUTHORS     Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and
             Weisshaar, B.
TITLE       High-throughput generation of sequence indexes from T-DNA
             mutagenized Arabidopsis thaliana lines
JOURNAL     Biotechniques 35 (6), 1164-1168 (2003)
PUBMED      14682050
REFERENCE   4 (bases 1 to 280)
AUTHORS     Strizhov, N., Li, Y., Rosso, M.G. and Weisshaar, B.
TITLE       Direct Submission
JOURNAL     Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
             This sequence has been recovered from the left border of the T-DNA.
             It indicates an insertion within the locus defined by BAC clone
             f3k23. Details on the protocols used for generation of the sequence
             are described in References 1-3. The sequences are generated at the
             MPI for Plant Breeding Research in the context of the GABI-Kat
             project. GABI-Kat is part of the German Plant Genomics program
             designated 'GABI'. Information on line availability can be found
             at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source             1..280
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /strain="Columbia 0"
                        /db_xref="taxon:3702"
                        /clone="GK-421D09-018141"
                        /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                        /ecotype="Col-0"
                        /notes="PCR was performed on DNA from Arabidopsis thaliana
                        plants (T1) which were transformed with the T-DNA from
                        vector pAC161 (GenBank accession number: AJ537514). The
                        lines contain one or more T-DNA insertions. The DNA
                        fragment(s) resulting from the PCR were directly sequenced
                        to determine the genomic sequence flanking the insertion.
                        T-DNA derived sequences were removed."
ORIGIN
Query Match      87.0%; Score 17.4; DB 9; Length 280;
Best Local Similarity 94.7%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db      353 GGCATGAACAAGAGATGATT 372

RESULT 6
BX288914/c
LOCUS      280 bp      DNA      linear      GSS 02-APR-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-421D09-018141,
             genomic survey sequence.
ACCESSION  BX288914
VERSION     BX288914.1 GI:28887910
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE   1
AUTHORS     Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weisshaar, B.
TITLE       GABI-Kat Simplesearch: a flanking sequence tag (fST) database for
             the identification of T-DNA insertion mutants in Arabidopsis
             thaliana
JOURNAL     Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE     22755829
PUBMED      12874060
REFERENCE   2
AUTHORS     Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and
             Weisshaar, B.
TITLE       An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
             flanking sequence tag-based reverse genetics
JOURNAL     Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE     23117147
PUBMED      14756321
REFERENCE   3
AUTHORS     Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and
             Weisshaar, B.
TITLE       High-throughput generation of sequence indexes from T-DNA
             mutagenized Arabidopsis thaliana lines
JOURNAL     Biotechniques 35 (6), 1164-1168 (2003)
PUBMED      14682050
REFERENCE   4 (bases 1 to 280)
AUTHORS     Strizhov, N., Li, Y., Rosso, M.G. and Weisshaar, B.
TITLE       Direct Submission
JOURNAL     Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
             This sequence has been recovered from the left border of the T-DNA.
             It indicates an insertion within the locus defined by BAC clone
             f3k23. Details on the protocols used for generation of the sequence
             are described in References 1-3. The sequences are generated at the
             MPI for Plant Breeding Research in the context of the GABI-Kat
             project. GABI-Kat is part of the German Plant Genomics program
             designated 'GABI'. Information on line availability can be found
             at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source             1..280
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /strain="Columbia 0"
                        /db_xref="taxon:3702"
                        /clone="GK-421D09-018141"
                        /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                        /ecotype="Col-0"
                        /notes="PCR was performed on DNA from Arabidopsis thaliana
                        plants (T1) which were transformed with the T-DNA from
                        vector pAC161 (GenBank accession number: AJ537514). The
                        lines contain one or more T-DNA insertions. The DNA
                        fragment(s) resulting from the PCR were directly sequenced
                        to determine the genomic sequence flanking the insertion.
                        T-DNA derived sequences were removed."
ORIGIN
Query Match      87.0%; Score 17.4; DB 9; Length 280;
Best Local Similarity 94.7%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db      353 GGCATGAACAAGAGATGATT 372

RESULT 7
BX288914/c
LOCUS      313 bp      mRNA      linear      EST 05-NOV-2003
DEFINITION Mus musculus cDNA clone NIA:A0504F09 IMAGE:30743204 5', mRNA
             sequence.
ACCESSION  CF906991
VERSION     CF906991.1 GI:38177928
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 313)
REFERENCE   1
AUTHORS     Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE       Construction of long-transcript enriched cDNA libraries from
             submicrogram amounts of total RNAs by a universal PCR amplification
             method
JOURNAL     Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE     21429098
PUBMED      11544199
COMMENT     Contact: Dawood B. Dudekula
             Laboratory of Genetics

Db      353 GGCATGAACAAGAGATGATT 372

RESULT 8
BX288914/c
LOCUS      305 bp      mRNA      linear      EST 15-NOV-2002
DEFINITION KS11017B12 KS11 Capsicum annuum cDNA, mRNA sequence.
ACCESSION  CA520799
VERSION     CA520799.1 GI:25034824
KEYWORDS    EST.
SOURCE      Capsicum annuum
ORGANISM    Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Capsicum.
            1 (bases 1 to 305)
REFERENCE   1
AUTHORS     Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
             Hur, C.-G. and Choi, D.
TITLE       Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
             annuum L.) and Sequence Analysis in Relation to Hypersensitive
             Response Against Pathogen
JOURNAL     Unpublished (2001)
COMMENT     Contact: Doil Choi
             Genome Research Center and National Center for Genome Information
             Korea Research Institute of Bioscience and Biotechnology
             P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
             Tel: 82-42-860-4340
             Fax: 82-42-860-4309
             Email: doil@mail.kribb.re.kr
             Plate: 017 row: B column: 12.
FEATURES             Location/Qualifiers
     source             1..305
                        /organism="Capsicum annuum"
                        /mol_type="mRNA"
                        /db_xref="taxon:4072"
                        /clone_lib="KS11"
ORIGIN
Query Match      87.0%; Score 17.4; DB 6; Length 305;
Best Local Similarity 94.7%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db      353 GGCATGAACAAGAGATGATT 372

RESULT 9
BX288914/c
LOCUS      313 bp      mRNA      linear      EST 05-NOV-2003
DEFINITION Mus musculus cDNA clone NIA:A0504F09 IMAGE:30743204 5', mRNA
             sequence.
ACCESSION  CF906991
VERSION     CF906991.1 GI:38177928
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 313)
REFERENCE   1
AUTHORS     Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE       Construction of long-transcript enriched cDNA libraries from
             submicrogram amounts of total RNAs by a universal PCR amplification
             method
JOURNAL     Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE     21429098
PUBMED      11544199
COMMENT     Contact: Dawood B. Dudekula
             Laboratory of Genetics

```

National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@gsun.grc.nia.nih.gov
Plate: A0504 row: F column: 09
Seq primer: M13 Reverse
High quality sequence stop: 313
POLYA-No.

FEATURES

source

Location/Qualifiers
1. 313
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H/He mice"
/db_xref="niaEST:A0504F09-5"
/db_xref="taxon:10090"
/clones="NTA:A0504F09 IMAGE:30743204"
/dev_stage="9-15C cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Akihiro Umezawa (Keio University School of Medicine, Japan). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]:
5'-pGACGATTTAGATCGCGCGCCCTTTT-3' from 2.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 Kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 313;
Best Local Similarity 94.7%; Pred. No. 1le+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATGAACAAGAGATGAT 19

Db 289 GAAATGAACAAGAGATGAT 271
|||||

RESULT 8

BX837806

LOCUS

DEFINITION BX837806 Arabidopsis thaliana Hormone Treated Callus Col-0
Arabidopsis thaliana cDNA clone GSLTPGH21ZC04 5PRIM, mRNA sequence.

ACCESSION

BX837806

VERSION

BX837806.1

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae;

Streptophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 423)

Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C.,

Menard, M., Craud, C., Quetier, F., Scarpelli, C., Schachter, V.,

Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

JOURNAL COMMENT

Annotation
Unpublished (2004)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallou O., Wincker P., Menard M., Craud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque Projet EF/EST
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source

Location/Qualifiers
1. 423
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTPGH21ZC04"
/tissue_type="Hormone Treated Callus"
/clone_lib="Arabidopsis thaliana Hormone Treated Callus Col-0"

ORIGIN

Query Match 87.0%; Score 17.4; DB 5; Length 423;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACATGAACAAGAGATGATT 20

Db 403 ACATGAACAAGAGATGATT 421
|||||

RESULT 9

AQ183243/c

LOCUS

DEFINITION

AQ183243 426 bp DNA linear GSS 01-NOV-1998
HS 3140 B2 B02 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3140 Col=4 Row=D, genomic survey sequence.

ACCESSION

AQ183243

VERSION

AQ183243.1

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 426)

Mahairas, G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,

Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and

Hood, L.

Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

99380589

10449764

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3140 row: D column: 4

Class: BAC ends

High quality sequence stop: 426.

Location/Qualifiers

```

source
1. .426
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=3140 Col=4 Row=D"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/notes="Organ: sperm; Vector: pBelBAC11; BAC Clones in E-Coli DH10B"

ORIGIN
Query Match      87.0%; Score 17.4; DB 8; Length 426;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACATGAACACAGAGATGATT 20
|||||
Db 273 ACATGAACACAGAGATGACT 255

RESULT 10
CN958899
LOCUS
DEFINITION
6399_100122_44 Fundulus heteroclitus Liver Fundulus heteroclitus
cDNA, mRNA sequence.
ACCESSION
CN958899
VERSION
CN958899.1 GI:48440488
KEYWORDS
EST.
SOURCE
Fundulus heteroclitus (killifish)
ORGANISM
Fundulus heteroclitus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Cyprinodontiformes; Fundulidae; Fundulus.
1 (bases 1 to 442)
Crawford,D.L., Oleksiak,M.F., Koleil,K.J., Paschall,J., VanWye,J.,
Roach,J.L. and Whitehead,J.A.
Fundulus Functional Genomics: EST Database for Teleost Fish
Unpublished (2004)
Contact: Crawford, Douglas L.
Marine Genomics - Crawford Lab
Rosenstiel School of Marine and Atmospheric Science - University of
Miami
4600 Rickenbacker Causeway, Miami, FL 33149-1098 USA
Tel: 305 361 4121
Email: dcrawford@rsmas.miami.edu
Database Web Interface
http://genomics.rsmas.miami.edu/funnybase/super_craw3/
Plate: 100122 row: E column: 6.
FEATURES
source
1. .442
/organism="Fundulus heteroclitus"
/mol_type="mRNA"
/db_xref="taxon:8078"
/tissue_type="Liver"
/clone_lib="Fundulus heteroclitus Liver"
/notes="Organ: Liver"

ORIGIN
Query Match      87.0%; Score 17.4; DB 7; Length 442;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATGAACACAGAGATGATT 19
|||||
Db 201 GTCATGACACAGAGATGAT 219

RESULT 11
AZ654527
LOCUS
DEFINITION
450 bp DNA linear GSS 14-DEC-2000
AZ654527
1M0528D16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0528D16 R, genomic survey sequence.

ACCESSION
AZ654527
VERSION
AZ654527.1 GI:11791673
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingley,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SUC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0528 row: D column: 16
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 450.
FEATURES
source
1. .450
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0528D16"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (GI4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match      87.0%; Score 17.4; DB 8; Length 450;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATGAACACAGAGATGATT 19
|||||
Db 132 GAAATGACACAGAGATGAT 150

RESULT 12
AK015232
LOCUS
DEFINITION
509 bp mRNA linear HTC 03-APR-2004
AK015232
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4930429C20 product:unclassified, full insert

```

sequence.
AK015232
AK015232.1 GI:12853489
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitzunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 509)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tegami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, I., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.jp/>) for further
details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to

sheared to 0.9-1 Kbp before ligation."

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 527;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACATGAACAAGAGATGATT 20
|||||
DB 160 ACATGAACAAGAGATGATT 142

RESULT 14

CF198522 570 bp mRNA linear EST 01-AUG-2003
LOCUS EST0117 Tamarix androssowii leaf Tamarix androssowii cDNA, mRNA
DEFINITION sequence.
ACCESSION CF198522
VERSION CF198522.1 GI:33392895
KEYWORDS EST.
SOURCE Tamarix androssowii
ORGANISM Tamarix androssowii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Tamaricaceae; Tamarix.

REFERENCE Wang, Y., Yang, C., Jiang, J., Liu, G., Wu, J. and Liu, Z.
AUTHORS EST acquired from cDNA library of Tamarix androssowii treated with
TITLE NaHCO3

JOURNAL

COMMENT Unpublished (2003)
Contact: Yucheng Wang
Forestry Source and Environment College
Northeast Forestry University
Hexing 26, Harbin, Heilongjiang, 150040, P.R. China
Tel: 086-451-2190607
Email: WANGYUCHENG1029@YAHOO.COM.CN.

FEATURES

source
1..570
/organism="Tamarix androssowii"
/mol_type="mRNA"
/db_xref="taxon:189785"
/tissue_type="leaf"
/clone_lib="Tamarix androssowii leaf"

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 570;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACATGAACAAGAGATGATT 20
|||||
DB 143 ACATGAACAAGAGGTGATT 161

RESULT 15

CF909462/c 574 bp mRNA linear EST 05-NOV-2003
LOCUS A0536H03-5 NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)
DEFINITION Mus musculus cDNA clone NIA:A0536H03 IMAGE:30746294 5', mRNA
sequence.

ACCESSION CF909462
VERSION CF909462.1 GI:38180399
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 574)
AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)

MEDLINE 21429098
PubMed 11544199
COMMENT

Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Casell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: A0536 row: H column: 03
Seq primer: M13 Reverse
High quality sequence stop: 574
POLYA=No.

FEATURES

Location/Qualifiers
source
1..574

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H/He mice"
/db_xref="niaEST:A0536H03-5"
/db_xref="taxon:10090"
/clone="NIA:A0536H03 IMAGE:30746294"
/dev_stage="9-15C cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library [Ref.
Genome Res. 11: 1553-1558 (2001). PMID: 11544199]. Total
RNAs were obtained from Dr. Akihiro Umezawa (Keio
University School of Medicine, Japan). Double-stranded
cDNAs were synthesized with an Oligo(dT) primer
[Invitrogen]:
5'-pGACTAGTTCATAGTCGAGCGCGCCCTTTT-3' from
2.2 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lone-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.5 kb. The library was
constructed by Yulan Piao."

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 574;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCATGAACAAGAGATGAT 19
|||
DB 510 GAAATGAACAAGAGATGAT 492

Search completed: March 17, 2005, 11:07:46
Job time : 1390.27 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-61
Perfect score: 20
Sequence: 1 gacgaagaacaagauguu 20
Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	30	6	AR027810 Sequence
2	20	100.0	30	6	AR027841 Sequence
C 3	20	100.0	87	6	AX151115 Sequence
C 4	20	100.0	99	14	HPBPRECA M76687 Hepatitis B
C 5	20	100.0	99	14	HPBPRECB M76688 Hepatitis B
C 6	20	100.0	99	14	HPBPRECC M76689 Hepatitis B
C 7	20	100.0	99	14	HPBPRECD M76690 Hepatitis B
C 8	20	100.0	99	14	HPBPRECE M76691 Hepatitis B
C 9	20	100.0	99	14	HPBPRECF M76692 Hepatitis B
C 10	20	100.0	99	14	HPBPRECG M76693 Hepatitis B
C 11	20	100.0	99	14	HPBPRECH M76694 Hepatitis B
C 12	20	100.0	99	14	HPBPRECI M76695 Hepatitis B
C 13	20	100.0	99	14	HPBPRECM M76699 Hepatitis B
C 14	20	100.0	129	6	AX151114 Sequence
C 15	20	100.0	150	14	AF528205 Hepatitis
C 16	20	100.0	150	14	AF528206 Hepatitis
C 17	20	100.0	150	14	AF528207 Hepatitis
C 18	20	100.0	150	14	AF528208 Hepatitis
C 19	20	100.0	150	14	AF528209 Hepatitis

C 20	100.0	150	14	AF528210 Hepatitis
C 21	100.0	150	14	AF528211 Hepatitis
C 22	100.0	150	14	AF528212 Hepatitis
C 23	100.0	150	14	AF528213 Hepatitis
C 24	100.0	150	14	AF528214 Hepatitis
C 25	100.0	150	14	AF528215 Hepatitis
C 26	100.0	150	14	AF528216 Hepatitis
C 27	100.0	150	14	AF528217 Hepatitis
C 28	100.0	150	14	AF528218 Hepatitis
C 29	100.0	150	14	AF528219 Hepatitis
C 30	100.0	150	14	AF528220 Hepatitis
C 31	100.0	150	14	AF528221 Hepatitis
C 32	100.0	150	14	AF528222 Hepatitis
C 33	100.0	150	14	AF528224 Hepatitis
C 34	100.0	150	14	AF528225 Hepatitis
C 35	100.0	150	14	AF528226 Hepatitis
C 36	100.0	150	14	AF528227 Hepatitis
C 37	100.0	150	14	AF528228 Hepatitis
C 38	100.0	150	14	AF528229 Hepatitis
C 39	100.0	150	14	AF528231 Hepatitis
C 40	100.0	150	14	AF528232 Hepatitis
C 41	100.0	150	14	AF528233 Hepatitis
C 42	100.0	150	14	AF528234 Hepatitis
C 43	100.0	150	14	AF528235 Hepatitis
C 44	100.0	150	14	AF528236 Hepatitis
C 45	100.0	150	14	AF528237 Hepatitis

ALIGNMENTS

RESULT 1
AR027810
LOCUS AR027810 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 8 from patent US 5856459.
ACCESSION AR027810
VERSION AR027810.1 GI:5938630
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and
Mills,J.S.
TITLE Oligonucleotides specific for Hepatitis B virus
Patent: US 5856459-A 8 05-JAN-1999;
FEATURES Location/Qualifiers
source
1..30
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAAAGAGAUU 20
||||:|||||:|:|:
Db 1 GACATCAACACAGATGATT 20
RESULT 2
AR027841
LOCUS AR027841 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 39 from patent US 5856459.
ACCESSION AR027841
VERSION AR027841.1 GI:5938661
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

```

Mills,J.S.
TITLE      Oligonucleotides specific for hepatitis B virus
JOURNAL    Patent: US 5856459-A 39 05-JAN-1999;
FEATURES   Location/Qualifiers
            source
            1..30
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACAUGAACAAAGAGAUU 20
      |||:|||||:|||||:|||||:
Db      1 GACATGACACAGAGATGATT 20

RESULT 3
AX151115/c      87 bp      DNA      linear      PAT 22-JUN-2001
LOCUS      AX151115
DEFINITION      Sequence 4 from Patent WO0138498.
ACCESSION      AX151115
VERSION        AX151115.1 GI:14533317
KEYWORDS       synthetic construct
SOURCE         synthetic construct
ORGANISM       other sequences; artificial sequences.
REFERENCE      1
AUTHORS        Scuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F.,
                Fried,M. and Roessau,R.
TITLE          A new genotype of hepatitis b virus
JOURNAL        Patent: WO 0138498-A 4 31-MAY-2001;
                Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES       Location/Qualifiers
            source
            1..87
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 87;
Best Local Similarity 80.0%; Pred. No. 51;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACAUGAACAAAGAGAUU 20
      |||:|||||:|||||:|||||:
Db      43 GACATGACACAGAGATGATT 24

RESULT 4
HPBPFECA/c      99 bp      DNA      linear      VRL 11-MAY-1994
LOCUS      HPBPFECA
DEFINITION      Hepatitis B virus type1 precore protein (pre-C region, C) gene, 5'
                end.
ACCESSION      M76687
VERSION        M76687.1 GI:485341
KEYWORDS       e antigen; precore protein; tolerogen.
SOURCE         Hepatitis B virus
ORGANISM       Hepatitis B virus
REFERENCE      1 (bases 1 to 99)
AUTHORS        Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
                Will,H.
TITLE          Prevalence and type of pre-C HBV mutants in anti-HBe positive
                carriers with chronic liver disease in a highly endemic area
JOURNAL        Virology 183 (2), 840-844 (1991)
MEDLINE        91306476
PUBMED         1853582
COMMENT        Original
FEATURES       Location/Qualifiers
            source
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"

variation
note="c in wt; t in virus type 2"
gene
10..93
/gene="C"
CDS
10..93
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precore protein"
/protein_id="AAA45508.1"
/db_xref="GI:485344"
/translation="WQLFHLCLIISCSCTVQASKLCIGWL"
/gene="C"
/note="g in wt; a in virus type 2 (creates internal stop
codon)"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACAUGAACAAAGAGAUU 20
      |||:|||||:|||||:|||||:
Db      52 GACATGACACAGAGATGATT 33

RESULT 5
HPBPFECA/c      99 bp      DNA      linear      VRL 11-MAY-1994
LOCUS      HPBPFECA
DEFINITION      Hepatitis B virus type 2precore protein (pre-C region, C) gene, 5'
                end.
ACCESSION      M76688
VERSION        M76688.1 GI:485343
KEYWORDS       e antigen; precore protein; tolerogen.
SOURCE         Hepatitis B virus
ORGANISM       Hepatitis B virus
REFERENCE      1 (bases 1 to 99)
AUTHORS        Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
                Will,H.
TITLE          Prevalence and type of pre-C HBV mutants in anti-HBe positive
                carriers with chronic liver disease in a highly endemic area
JOURNAL        Virology 183 (2), 840-844 (1991)
MEDLINE        91306476
PUBMED         1853582
COMMENT        Original
FEATURES       Location/Qualifiers
            source
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"

variation
note="c in wt; t in virus type 2"
gene
10..93
/gene="C"
CDS
10..93
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precore protein"
/protein_id="AAA45508.1"
/db_xref="GI:485344"
/translation="WQLFHLCLIISCSCTVQASKLCIGWL"
/gene="C"
/note="g in wt; a in virus type 2 (creates internal stop
codon)"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

/mol_type="genomic DNA"
/db_xref="taxon:10407"
10..93
/gene="C"
10..93
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precore protein"
/protein_id="AAA45507.1"
/db_xref="GI:485342"
/translation="WQLFHLCLIISCSCTVQASKLCIGWL"
92
/gene="C"
/note="g in wt; a in virus type 1 (creates internal stop
codon)"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 GACAUGAACAAAGAGAUU 20
      |||:|||||:|||||:|||||:
Db      52 GACATGACACAGAGATGATT 33

RESULT 5
HPBPFECA/c      99 bp      DNA      linear      VRL 11-MAY-1994
LOCUS      HPBPFECA
DEFINITION      Hepatitis B virus type 2precore protein (pre-C region, C) gene, 5'
                end.
ACCESSION      M76688
VERSION        M76688.1 GI:485343
KEYWORDS       e antigen; precore protein; tolerogen.
SOURCE         Hepatitis B virus
ORGANISM       Hepatitis B virus
REFERENCE      1 (bases 1 to 99)
AUTHORS        Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and
                Will,H.
TITLE          Prevalence and type of pre-C HBV mutants in anti-HBe positive
                carriers with chronic liver disease in a highly endemic area
JOURNAL        Virology 183 (2), 840-844 (1991)
MEDLINE        91306476
PUBMED         1853582
COMMENT        Original
FEATURES       Location/Qualifiers
            source
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"

variation
note="c in wt; t in virus type 2"
gene
10..93
/gene="C"
CDS
10..93
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precore protein"
/protein_id="AAA45508.1"
/db_xref="GI:485344"
/translation="WQLFHLCLIISCSCTVQASKLCIGWL"
92
/gene="C"
/note="g in wt; a in virus type 2 (creates internal stop
codon)"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GACAUGAACAGAGAUU 20
 ||||:|||||||:|:|:
 Db 52 GACATGAACAAGATGATT 33

RESULT 6
 HPBP/CC/c
 LOCUS Hepatitis B virus type 3precure protein (pre-C region, C) gene, 5'
 DEFINITION end.

ACCESSION M76689
 VERSION M76691.1 GI:485345
 KEYWORDS e antigen; precure protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

REFERENCE 1 (bases 1 to 99)
 AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582

COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES Location/Qualifiers
 source 1..99
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10407"
 gene 10..93
 /gene="C"
 CDS 10..93
 /gene="C"
 /standard_name="pre-C region"
 /codon_start=1
 /product="precure protein"
 /protein_id="AAA45510.1"
 /db_xref="GI:485348"
 /translation="MQLPHLCIIISCSPTFOASKLCGLWL"
 variation 92
 /gene="C"
 /note="g in wt; a in virus type 4 (creates internal stop codon)"

ORIGIN
 Location/Qualifiers
 source 1..99
 /note="g in wt; a in virus type 4 (gly to asp)"

variation 95
 /note="g in wt; a in virus type 4 (gly to asp)"

Query Match 100.0%; Score 20; DB 14; Length 99;
 Best Local Similarity 80.0%; Pred. No. 50;
 Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUU 20
 ||||:|||||||:|:|:
 Db 52 GACATGAACAAGATGATT 33

RESULT 8
 HPBP/CC/c
 LOCUS Hepatitis B virus type 5 precure protein (pre-C region, C) gene, 5'
 DEFINITION end.

ACCESSION M76691
 VERSION M76691.1 GI:485349
 KEYWORDS e antigen; precure protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

REFERENCE 1 (bases 1 to 99)
 AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582

COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES Location/Qualifiers
 source 1..99
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10407"
 gene 10..93
 /gene="C"
 CDS 10..93
 /gene="C"
 /standard_name="pre-C region"
 /codon_start=1
 /product="precure protein"
 /protein_id="AAA45511.1"
 /db_xref="GI:485348"
 /translation="MQLPHLCIIISCSPTFOASKLCGLWL"

ORIGIN
 Location/Qualifiers
 source 1..99
 /note="g in wt; t in virus type 3 (val to phe)"

variation 92
 /note="g in wt; t in virus type 3 (val to phe)"

Query Match 100.0%; Score 20; DB 14; Length 99;
 Best Local Similarity 80.0%; Pred. No. 50;
 Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUU 20
 ||||:|||||||:|:|:
 Db 52 GACATGAACAAGATGATT 33

RESULT 7
 HPBP/CC/c
 LOCUS Hepatitis B virus type 4 precure protein (pre-C region, C) gene, 5'
 DEFINITION end.

ACCESSION M76690
 VERSION M76690.1 GI:485347
 KEYWORDS e antigen; precure protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

REFERENCE 1 (bases 1 to 99)

AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582

COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES Location/Qualifiers
 source 1..99
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10407"
 gene 10..93
 /gene="C"
 CDS 10..93
 /gene="C"
 /standard_name="pre-C region"
 /codon_start=1
 /product="precure protein"
 /protein_id="AAA45510.1"
 /db_xref="GI:485348"
 /translation="MQLPHLCIIISCSPTFOASKLCGLWL"
 variation 92
 /gene="C"
 /note="g in wt; a in virus type 4 (creates internal stop codon)"

ORIGIN
 Location/Qualifiers
 source 1..99
 /note="g in wt; a in virus type 4 (gly to asp)"

variation 95
 /note="g in wt; a in virus type 4 (gly to asp)"

Query Match 100.0%; Score 20; DB 14; Length 99;
 Best Local Similarity 80.0%; Pred. No. 50;
 Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAUU 20
 ||||:|||||||:|:|:
 Db 52 GACATGAACAAGATGATT 33

RESULT 8
 HPBP/CC/c
 LOCUS Hepatitis B virus type 5 precure protein (pre-C region, C) gene, 5'
 DEFINITION end.

ACCESSION M76691
 VERSION M76691.1 GI:485349
 KEYWORDS e antigen; precure protein; tolerogen.
 SOURCE Hepatitis B virus
 ORGANISM Hepatitis B virus
 Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

REFERENCE 1 (bases 1 to 99)
 AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
 TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
 JOURNAL Virology 183 (2), 840-844 (1991)
 MEDLINE 91306476
 PUBMED 1853582

COMMENT Original source text: Hepatitis B virus DNA.
 FEATURES Location/Qualifiers
 source 1..99
 /organism="Hepatitis B virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10407"
 gene 10..93
 /gene="C"
 CDS 10..93
 /gene="C"
 /standard_name="pre-C region"
 /codon_start=1
 /product="precure protein"
 /protein_id="AAA45511.1"
 /db_xref="GI:485348"
 /translation="MQLPHLCIIISCSPTFOASKLCGLWL"

```

/db xref="GI:485350"
/translation="MQFLHCLIIISCSPTVQPSKLCIGWL"
64..67
/gene="C"
/note="gcc in wt; ccg in virus type 5 (ala to pro)"
92
/gene="C"
/note="g in wt; a in virus type 5 (creates internal stop codon)"
95
ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUU 20
    |||:|||||:|||||:|:|:|
Db 52 GACATGAACAAGAGATGATT 33

RESULT 9
HPBPREFC/c
LOCUS
DEFINITION
Hepatitis B virus type 6 precore protein (pre-C region, C) gene, 5' end.
ACCESSION M76692
VERSION M76692.1 GI:485351
KEYWORDS e antigen; precore protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
    source
        Location/Qualifiers
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            10..93
            /gene="C"
            10..93
            misc_feature
            /gene="C"
            /product="precure protein"
            /standard_name="pre-C region note: putative CDS"
            10
            variation
            /note="a in wt; t in virus type 7 (loss of start codon)"
            14
            /gene="C"
            /note="a in wt; g in virus type 7 (gln to arg)"
            92
            /gene="C"
            /note="g in wt; a in virus type 7 (creates internal stop codon)"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUU 20
    |||:|||||:|||||:|:|:|
Db 52 GACATGAACAAGAGATGATT 33

RESULT 11
HPBPREFC/c
LOCUS
DEFINITION
Hepatitis B virus type 8 precore protein (pre-C region, C) gene, 5' end.
ACCESSION M76694
VERSION M76694.1 GI:485353
KEYWORDS e antigen; precore protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
    source
        Location/Qualifiers
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            10..93
            /gene="C"
            10..93
            misc_feature
            /gene="C"
            /product="precure protein"
            /notes="putative cds"
            11
            variation
            /gene="C"
            /note="t in wt; c in virus type 6 (loss of start codon)"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUU 20
    |||:|||||:|||||:|:|:|
Db 52 GACATGAACAAGAGATGATT 33

RESULT 10
HPBPREFC/c
LOCUS
DEFINITION
Hepatitis B virus type 7 precore protein (pre-C region, C) gene, 5' end.
ACCESSION M76693
VERSION M76693.1 GI:485352

```

```

KEYWORDS e antigen; precore protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
    source
        Location/Qualifiers
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            10..93
            /gene="C"
            10..93
            misc_feature
            /gene="C"
            /product="precure protein"
            /standard_name="pre-C region note: putative CDS"
            10
            variation
            /note="a in wt; t in virus type 7 (loss of start codon)"
            14
            /gene="C"
            /note="a in wt; g in virus type 7 (gln to arg)"
            92
            /gene="C"
            /note="g in wt; a in virus type 7 (creates internal stop codon)"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUU 20
    |||:|||||:|||||:|:|:|
Db 52 GACATGAACAAGAGATGATT 33

RESULT 11
HPBPREFC/c
LOCUS
DEFINITION
Hepatitis B virus type 8 precore protein (pre-C region, C) gene, 5' end.
ACCESSION M76694
VERSION M76694.1 GI:485353
KEYWORDS e antigen; precore protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
    source
        Location/Qualifiers
            1..99
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            10..93
            /gene="C"
            10..93
            misc_feature
            /gene="C"

```

```
/gene="C"
/product="precure protein"
12
/standard_name="pre-C region note: putative CDS"
variation
12
/gene="C"
/notes="g in wt; t in virus type 8 (loss of start codon)"
92
/gene="C"
/notes="g in wt; a in virus type 8 (creates internal stop codon)"
95
/notes="g in wt; a in virus type 8 (gly to asp)"
ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGAGAUU 20
||||:|||||
Db 52 GACATGAACAAGAGATGATT 33

RESULT 12
HPBPREC1/c
LOCUS
DEFINITION
Hepatitis B virus type 9 precure protein (pre-C region, C) gene, 5' end.
ACCESSION M76695.1 GI:485354
VERSION
KEYWORDS e antigen; precure protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
source
1..99
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:10407"
10..99
/gene="C"
10..99
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precure protein"
/protein_id="AAA4515.1"
/db_xref="GI:485362"
/translation="MOLFHLIIISCSPTVQASKLCLGLWDM"
95
variation
/gene="C"
/notes="g in wt; a in virus type 13 (gly to asp)"
ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGAGAUU 20
||||:|||||
Db 52 GACATGAACAAGAGATGATT 33

RESULT 14
AX151114/c
LOCUS
DEFINITION
Sequence 3 from Patent WO0138498.
ACCESSION AX151114
VERSION AX151114.1 GI:14533316
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F., Fried,M. and Rossau,R.
TITLE A new genotype of hepatitis b virus
JOURNAL Patent: WO 0138498-A 3 31-MAY-2001; Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES
source
1..129
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGAGAUU 20
||||:|||||
Db 52 GACATGAACAAGAGATGATT 33
```

```
HPBPREC1/c
LOCUS
DEFINITION
Hepatitis B virus type 13 precure protein (pre-C region, C) gene, 5' end.
ACCESSION M76699
VERSION M76699.1 GI:485361
KEYWORDS e antigen; precure protein; tolerogen.
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1 (bases 1 to 99)
AUTHORS Santantonio,T., Jung,M.C., Miska,S., Pastore,G., Pape,G.R. and Will,H.
TITLE Prevalence and type of pre-C HBV mutants in anti-HBe positive carriers with chronic liver disease in a highly endemic area
JOURNAL Virology 183 (2), 840-844 (1991)
MEDLINE 91306476
PUBMED 1853582
COMMENT Original source text: Hepatitis B virus DNA.
FEATURES
source
1..99
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:10407"
10..99
/gene="C"
10..99
/gene="C"
/standard_name="pre-C region"
/codon_start=1
/product="precure protein"
/protein_id="AAA4515.1"
/db_xref="GI:485362"
/translation="MOLFHLIIISCSPTVQASKLCLGLWDM"
95
variation
/gene="C"
/notes="g in wt; a in virus type 13 (gly to asp)"
ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 99;
Best Local Similarity 80.0%; Pred. No. 50;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAGAGAGAUU 20
||||:|||||
Db 52 GACATGAACAAGAGATGATT 33

RESULT 14
AX151114/c
LOCUS
DEFINITION
Sequence 3 from Patent WO0138498.
ACCESSION AX151114
VERSION AX151114.1 GI:14533316
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Stuyver,L., Schinazi,R., de Gendt,S., van Geyt,C., Zoulim,F., Fried,M. and Rossau,R.
TITLE A new genotype of hepatitis b virus
JOURNAL Patent: WO 0138498-A 3 31-MAY-2001; Pharmasset, Inc. (US) ; INNOGENETICS N.V. (BE)
FEATURES
source
1..129
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN
```

```

Query Match      100.0%; Score 20; DB 6; Length 129;
Best Local Similarity 80.0%; Pred. No. 49;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GACAUGAACAAAGAGAUAU 20
        |||:|||||:|||||:|:|:
Db      43 GACATGAACAAGAGATGATT 24

RESULT 15
AF528205/c
LOCUS      AF528205      150 bp      DNA      linear      VRL 31-JUL-2003
DEFINITION Hepatitis B virus ASC1123 core antigen precursor, gene, partial cds.
ACCESSION  AF528205
VERSION    AF528205.1 GI:32810971
KEYWORDS
SOURCE
ORGANISM   Hepatitis B virus
            Hepatitis B virus
            Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE  1 (bases 1 to 150)
            Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
            Comparative evaluation of HBV precore and basal core promoter
            mutants in Indian patients with diverse clinical manifestations
            Unpublished
REFERENCE  2 (bases 1 to 150)
            Gandhe,S.S., Chadha,M.S., Walimbe,A.M. and Arankalle,V.A.
            Direct Submission
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..150
/organism="Hepatitis B virus"
/proviral
/mol_type="genomic DNA"
/isolate="ASC1123"
/isolation_source="asymptomatic HBsAg carrier"
/specific_host="Homo sapiens"
/db_xref="taxon:10407"
/country="India"
<1..>150
/notes="contains partial basal core promoter"
64..>150
/notes="contains complete precore region"
/codon_start=1
/product="core antigen precursor"
/protein_id="AAP87556.1"
/db_xref="GI:32810972"
/translation="MQLFHLCLIISCSPTVQASKLCLGLXG"

misc_feature
CDS

ORIGIN

Query Match      100.0%; Score 20; DB 14; Length 150;
Best Local Similarity 80.0%; Pred. No. 48;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GACAUGAACAAAGAGAUAU 20
        |||:|||||:|||||:|:|:
Db      106 GACATGAACAAGAGATGATT 87

```

Search completed: March 17, 2005, 08:14:16
 Job time : 683.733 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-61
Perfect score: 20
Sequence: 1 gacgaagacaagagaauu 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	2 AAT72566	Aat72566 Hepatitis
2	20	100.0	20	2 AAT72565	Aat72565 Hepatitis
3	20	100.0	30	2 AAT72562	Aat72562 Hepatitis
4	20	100.0	30	2 AAT72563	Aat72563 Hepatitis
5	20	100.0	30	2 AAT72616	Aat72616 Hepatitis
6	20	100.0	30	2 AAT72617	Aat72617 Hepatitis
7	20	100.0	39	10 ADC64742	Adc64742 Hepatitis
8	20	100.0	87	4 AAD09094	Aad09094 Hepatitis
9	20	100.0	129	4 AAD09093	Aad09093 Hepatitis
10	20	100.0	639	6 AAD27422	Aad27422 Hepatitis
11	20	100.0	639	6 AAD31509	Aad31509 Hepatitis
12	20	100.0	655	4 AAH77569	Aah77569 HBV genot
13	20	100.0	655	4 AAH77568	Aah77568 HBV genot
14	20	100.0	655	4 AAH77574	Aah77574 HBV genot
15	20	100.0	655	4 AAH77573	Aah77573 HBV genot
16	20	100.0	655	4 AAH77570	Aah77570 HBV genot
17	20	100.0	655	4 AAH77571	Aah77571 HBV genot
18	20	100.0	664	4 AAH77572	Aah77572 HBV genot
19	20	100.0	669	12 ADO07220	Ado07220 Hepatitis
20	20	100.0	673	4 AAD09092	Aad09092 Hepatitis

C 21	20	100.0	675	4 AAH77563	Aah77563 HBV preCo
C 22	20	100.0	681	4 AAH77567	Aah77567 HBV genot
C 23	20	100.0	1395	2 AAV82688	Aav82688 Fulminant
C 24	20	100.0	1400	2 AAV82687	Aav82687 Fulminant
C 25	20	100.0	1445	2 AAV82692	Aav82692 Fulminant
C 26	20	100.0	1445	2 AAV82685	Aav82685 Fulminant
C 27	20	100.0	1445	2 AAV82690	Aav82690 Fulminant
C 28	20	100.0	1445	2 AAV82684	Aav82684 Fulminant
C 29	20	100.0	1500	2 AAV82695	Aav82695 Fulminant
C 30	20	100.0	1500	2 AAV82683	Aav82683 Fulminant
C 31	20	100.0	1500	2 AAV82694	Aav82694 Fulminant
C 32	20	100.0	1500	2 AAV82686	Aav82686 Fulminant
C 33	20	100.0	1500	2 AAV82706	Aav82706 Wild type
C 34	20	100.0	1500	2 AAV82689	Aav82689 Fulminant
C 35	20	100.0	1500	2 AAV82693	Aav82693 Fulminant
C 36	20	100.0	2342	1 AAN93072	Aan93072 Sequence
C 37	20	100.0	2743	1 AAN00003	Aan00003 Sequence
C 38	20	100.0	2743	2 AAQ04799	Aaq04799 Recombina
C 39	20	100.0	3180	4 AAH42375	Aah42375 Nucleotid
C 40	20	100.0	3182	6 AAD31765	Aad31765 Hepatitis
C 41	20	100.0	3182	9 ACA62422	Acac62422 Hepatitis
C 42	20	100.0	3182	10 AAD60866	Aad60866 Hepatitis
C 43	20	100.0	3220	3 AA288924	Aaz88924 Hepatitis
C 44	20	100.0	3248	4 AAD09091	Aad09091 Hepatitis
C 45	20	100.0	3248	4 AAH77562	Aah77562 HBV genot

ALIGNMENTS

RESULT 1
AAT72566
ID AAT72566 standard; RNA; 20 BP.
XX
AC AAT72566;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV46MYb.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..20
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
FT modified_base 1
FT /*tag= b
FT /mod_base= gm
FT modified_base 2
FT /*tag= c
FT /mod_base= OTHER
FT modified_base 3
FT /note= "2'-O-methyladenosine"
FT modified_base 4
FT /tag= d
FT /mod_base= cm
FT modified_base 5
FT /*tag= e
FT /mod_base= OTHER
FT modified_base 6
FT /tag= f
FT /mod_base= um
FT modified_base 7
FT /tag= g
FT /mod_base= gm
FT modified_base 8
FT /*tag= h
FT /mod_base= OTHER
FT modified_base 9
FT /note= "2'-O-methyladenosine"

```
FT FT /tag= i
FT FT /mod_base= OTHER
FT FT /note= "2'-O-methyladenosine"
FT modified_base
FT /tag= j
FT /mod_base= cm
FT modified_base
FT /tag= k
FT /mod_base= OTHER
FT FT /note= "2'-O-methyladenosine"
FT modified_base
FT /tag= l
FT /mod_base= OTHER
FT FT /note= "2'-O-methyladenosine"
FT modified_base
FT /tag= m
FT /mod_base= gm
FT modified_base
FT /tag= n
FT /mod_base= OTHER
FT FT /note= "2'-O-methyladenosine"
FT modified_base
FT /tag= o
FT /mod_base= gm
FT modified_base
FT /tag= p
FT /mod_base= OTHER
FT FT /note= "2'-O-methyladenosine"
FT modified_base
FT /tag= q
FT /mod_base= um
FT modified_base
FT /tag= r
FT /mod_base= gm
FT modified_base
FT /tag= s
FT /mod_base= OTHER
FT FT /note= "2'-O-methyladenosine"
FT modified_base
FT /tag= t
FT /mod_base= um
FT modified_base
FT /tag= u
FT /mod_base= um
FT
FT
XX PN WO9639502-A1.
XX XX
XX PD 12-DEC-1996.
XX XX
XX PF 04-JUN-1996; 96WO-EP002432.
XX PF 06-JUN-1995; 95US-00467397.
XX PR
XX PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX PA (HYBR-) HYBRIDON INC.
XX PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
XX PI Roberts NA, Roberts PC, Slade A;
XX DR WPI; 1997-043124/04.
XX XX
XX PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
XX PT used in the detection and treatment of HBV infection.
XX PS Claim 1; Page 12; 81pp; English.
XX CC The present sequence represents a synthetic oligonucleotide HBV46MYB
XX CC which is complementary to a portion of the hepatitis B virus (HBV) RNA.
XX CC The antisense oligonucleotide may be used to detect the presence of HBV
XX CC in a sample. The antisense oligonucleotide, and oligonucleotides
XX CC containing a sequence which is complementary to at least two non-
XX CC contiguous regions of an HBV nucleic acid, may be used for inhibiting HBV
XX CC replication in a cell or for the treatment of HBV infection
```

```
XX SQ Sequence 20 BP; 9 A; 2 C; 5 G; 0 T; 4 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 9,6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGAUU 20
Db 1 GACAUGAACAGAGAGAUU 20

RESULT 2
AAT72565
ID AAT72565 standard; DNA; 20 BP.
XX AC AAT72565;
XX DT 03-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV46YB.
XX KW HBV; HBV infection; inhibition; replication; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT misc_feature 1..20
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
XX PN WO9639502-A1.
XX PD 12-DEC-1996.
XX PF 04-JUN-1996; 96WO-EP002432.
XX PR 06-JUN-1995; 95US-00467397.
XX PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX PA (HYBR-) HYBRIDON INC.
XX PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
XX PI Roberts NA, Roberts PC, Slade A;
XX DR WPI; 1997-043124/04.
XX XX
XX PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
XX PT used in the detection and treatment of HBV infection.
XX PS Claim 1; Page 12; 81pp; English.
XX CC The present sequence represents a synthetic oligonucleotide HBV46YB which
XX CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
XX CC antisense oligonucleotide may be used to detect the presence of HBV in a
XX CC sample. The antisense oligonucleotide, and oligonucleotides containing a
XX CC sequence which is complementary to at least two non- contiguous regions
XX CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
XX CC cell or for the treatment of HBV infection
XX SQ Sequence 20 BP; 9 A; 2 C; 5 G; 4 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 9,6;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGAUU 20
Db 1 GACATGAACAAGAGATGATT 20

RESULT 3
AAT72562
```

```

ID AAT72562 standard; DNA; 30 BP.
XX AC
XX AAT72562;
XX DT
XX 03-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV88b.
XX KW HBV; HBV infection; inhibition; replication; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT misc_feature 1..30
XX FT /tag= a
XX FT /note= "Internucleotide linkages are phosphorothioate"
XX PN
XX WO9639502-A1.
XX PD 12-DEC-1996.
XX PF 04-JUN-1996; 96WO-EP002432.
XX PR 06-JUN-1995; 95US-00467397.
XX PA (HOFF ) HOPFMANN LA ROCHE & CO AG F.
XX PA (HYBR-) HYBRIDON INC.
XX PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
XX PI Roberts NA, Roberts PC, Siade A;
XX DR WPI; 1997-043124/04.
XX PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
XX PT used in the detection and treatment of HBV infection.
XX PS Claim 1; Page 12; 81pp; English.
XX CC The present sequence represents a synthetic oligonucleotide HBV88b which
XX CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
XX CC antisense oligonucleotide may be used to detect the presence of HBV in a
XX CC sample. The antisense oligonucleotide, and oligonucleotides containing a
XX CC sequence which is complementary to at least two non- contiguous regions
XX CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
XX CC cell or for the treatment of HBV infection
XX SQ Sequence 30 BP; 12 A; 3 C; 10 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 80.0%; Pred. NO. 9.9;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGUUGUU 20
DB 1 GACATGAACAGAGAGATGATT 20

RESULT 4
AAT72563
ID AAT72563 standard; DNA; 30 BP.
XX AC
XX AAT72563;
XX DT
XX 03-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV88Mb.
XX KW HBV; HBV infection; inhibition; replication; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT misc_feature 1..30

```

```

FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
FT 1..20
FT /tag= b
FT /note= "2'-Ome RNA"
FT modified_base 1
FT /tag= c
FT /mod_base= gm
FT modified_base 2
FT /tag= d
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 3
FT /tag= e
FT /mod_base= cm
FT modified_base 4
FT /tag= f
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 5
FT /tag= g
FT /mod_base= um
FT modified_base 6
FT /tag= h
FT /mod_base= gm
FT modified_base 7
FT /tag= i
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 8
FT /tag= j
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 9
FT /tag= k
FT /mod_base= cm
FT modified_base 10
FT /tag= l
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 11
FT /tag= m
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 12
FT /tag= n
FT /mod_base= gm
FT modified_base 13
FT /tag= o
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 14
FT /tag= p
FT /mod_base= gm
FT modified_base 15
FT /tag= q
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 16
FT /tag= r
FT /mod_base= um
FT modified_base 17
FT /tag= s
FT /mod_base= gm
FT modified_base 18
FT /tag= s
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 19
FT /tag= u
FT /mod_base= um
FT modified_base 20
FT /tag= v

```

PI Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
XX Claim 5; Page 15; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV-89b which
CC contains a sequence which is complementary to at least two non-contiguous
CC regions of a Hepatitis B virus (HBV) nucleic acid. The antisense
CC oligonucleotide may be used to detect the presence of HBV in a sample.
CC The antisense oligonucleotide, and oligonucleotides complementary to a
CC portion of the HBV RNA, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection
XX
SQ Sequence 30 BP; 12 A; 3 C; 9 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 80.0%; Pred. No. 9.9;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUAU 20
DB |||||:|||||:|||||:
1 GACATGACACAGAGATGATT 20

RESULT 6
AAT72617
ID AAT72617 standard; DNA; 30 BP.
XX AC AAT72617;
XX DT 04-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV-89Mb.
XX KW HBV; HBV infection; inhibition; replication; ss.
XX OS Synthetic.

FH Key Location/Qualifiers
FT misc_feature 1..30 /tag= a /note= "Internucleotide linkages are phosphorothioate"
FT misc_RNA 1..20 /tag= b /note= "2'-OMe RNA"
FT modified_base 1 /tag= c /mod_base= gm
FT modified_base 2 /tag= d /mod_base= OTHER
FT modified_base 3 /note= "2'-O-methyladenosine"
FT modified_base 4 /tag= e /mod_base= cm
FT modified_base 5 /tag= f /mod_base= OTHER
FT modified_base 6 /note= "2'-O-methyladenosine"
FT modified_base 7 /tag= g /mod_base= um
FT modified_base 8 /tag= h /mod_base= gm
FT modified_base 9 /tag= i /mod_base= OTHER
FT modified_base 10 /note= "2'-O-methyladenosine"

FT /mod_base= um

WO9639502-A1.
12-DEC-1996.
04-JUN-1996; 96WO-EF002432.
06-JUN-1995; 95US-00467397.
(HOFF) HOFFMANN LA ROCHE & CO AG F.
(HYBR-) HYBRIDON INC.
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
XX Claim 1; Page 12; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV88Mb which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a
CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non-contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection
XX
SQ Sequence 30 BP; 12 A; 3 C; 10 G; 1 T; 4 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUAU 20
DB |||||:|||||:|||||:
1 GACAUGAACACAGAGAUAU 20

RESULT 5
AAT72616
ID AAT72616 standard; DNA; 30 BP.
XX AC AAT72616;
XX DT 04-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV-89b.
XX KW HBV; HBV infection; inhibition; replication; ss.
XX OS Synthetic.

FH Key Location/Qualifiers
FT misc_feature 1..30 /tag= a /note= "Internucleotide linkages are phosphorothioate"
FT modified_base 1 /tag= b /mod_base= gm
FT modified_base 2 /tag= c /mod_base= OTHER
FT modified_base 3 /note= "2'-O-methyladenosine"
FT modified_base 4 /tag= d /mod_base= cm
FT modified_base 5 /tag= e /mod_base= OTHER
FT modified_base 6 /note= "2'-O-methyladenosine"
FT modified_base 7 /tag= f /mod_base= um
FT modified_base 8 /tag= g /mod_base= gm
FT modified_base 9 /tag= h /mod_base= OTHER
FT modified_base 10 /note= "2'-O-methyladenosine"

WO9639502-A1.
12-DEC-1996.
04-JUN-1996; 96WO-EF002432.
06-JUN-1995; 95US-00467397.
(HOFF) HOFFMANN LA ROCHE & CO AG F.
(HYBR-) HYBRIDON INC.
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
XX Claim 1; Page 12; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV88Mb which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a
CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non-contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection
XX
SQ Sequence 30 BP; 12 A; 3 C; 10 G; 1 T; 4 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUAU 20
DB |||||:|||||:|||||:
1 GACAUGAACACAGAGAUAU 20

RESULT 5
AAT72616
ID AAT72616 standard; DNA; 30 BP.
XX AC AAT72616;
XX DT 04-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV-89b.
XX KW HBV; HBV infection; inhibition; replication; ss.
XX OS Synthetic.

FH Key Location/Qualifiers
FT misc_feature 1..30 /tag= a /note= "Internucleotide linkages are phosphorothioate"
FT modified_base 1 /tag= b /mod_base= gm
FT modified_base 2 /tag= c /mod_base= OTHER
FT modified_base 3 /note= "2'-O-methyladenosine"
FT modified_base 4 /tag= d /mod_base= cm
FT modified_base 5 /tag= e /mod_base= OTHER
FT modified_base 6 /note= "2'-O-methyladenosine"
FT modified_base 7 /tag= f /mod_base= um
FT modified_base 8 /tag= g /mod_base= gm
FT modified_base 9 /tag= h /mod_base= OTHER
FT modified_base 10 /note= "2'-O-methyladenosine"

WO9639502-A1.
12-DEC-1996.
04-JUN-1996; 96WO-EF002432.
06-JUN-1995; 95US-00467397.
(HOFF) HOFFMANN LA ROCHE & CO AG F.
(HYBR-) HYBRIDON INC.
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
XX Claim 1; Page 12; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV88Mb which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a
CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non-contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection
XX
SQ Sequence 30 BP; 12 A; 3 C; 10 G; 1 T; 4 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUAU 20
DB |||||:|||||:|||||:
1 GACAUGAACACAGAGAUAU 20

RESULT 5
AAT72616
ID AAT72616 standard; DNA; 30 BP.
XX AC AAT72616;
XX DT 04-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV-89b.
XX KW HBV; HBV infection; inhibition; replication; ss.
XX OS Synthetic.

FH Key Location/Qualifiers
FT misc_feature 1..30 /tag= a /note= "Internucleotide linkages are phosphorothioate"
FT modified_base 1 /tag= b /mod_base= gm
FT modified_base 2 /tag= c /mod_base= OTHER
FT modified_base 3 /note= "2'-O-methyladenosine"
FT modified_base 4 /tag= d /mod_base= cm
FT modified_base 5 /tag= e /mod_base= OTHER
FT modified_base 6 /note= "2'-O-methyladenosine"
FT modified_base 7 /tag= f /mod_base= um
FT modified_base 8 /tag= g /mod_base= gm
FT modified_base 9 /tag= h /mod_base= OTHER
FT modified_base 10 /note= "2'-O-methyladenosine"

WO9639502-A1.
12-DEC-1996.
04-JUN-1996; 96WO-EF002432.
06-JUN-1995; 95US-00467397.
(HOFF) HOFFMANN LA ROCHE & CO AG F.
(HYBR-) HYBRIDON INC.
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
XX Claim 1; Page 12; 81pp; English.
XX
XX The present sequence represents a synthetic oligonucleotide HBV88Mb which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a
CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non-contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection
XX
SQ Sequence 30 BP; 12 A; 3 C; 10 G; 1 T; 4 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACACAGAGAUAU 20
DB |||||:|||||:|||||:
1 GACAUGAACACAGAGAUAU 20

RESULT 5
AAT72616
ID AAT72616 standard; DNA; 30 BP.
XX AC AAT72616;
XX DT 04-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV-89b.
XX KW HBV; HBV infection; inhibition; replication; ss.
XX OS Synthetic.

FH Key Location/Qualifiers
FT misc_feature 1..30 /tag= a /note= "Internucleotide linkages are phosphorothioate"
FT modified_base 1 /tag= b /mod_base= gm
FT modified_base 2 /tag= c /mod_base= OTHER
FT modified_base 3 /note= "2'-O-methyladenosine"
FT modified_base 4 /tag= d /mod_base= cm
FT modified_base 5 /tag= e /mod_base=

```
FT modified_base 8 /*tag= j
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 9
FT /*tag= k
FT /mod_base= cm
FT modified_base 10
FT /*tag= l
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 11
FT /*tag= m
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 12
FT /*tag= n
FT /mod_base= gm
FT modified_base 13
FT /*tag= o
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 14
FT /*tag= p
FT /mod_base= gm
FT modified_base 15
FT /*tag= q
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 16
FT /*tag= r
FT /mod_base= um
FT modified_base 17
FT /*tag= s
FT /mod_base= gm
FT modified_base 18
FT /*tag= t
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"
FT modified_base 19
FT /*tag= u
FT /mod_base= um
FT modified_base 20
FT /*tag= v
FT /mod_base= um
FT
FT WO9639502-A1.
FT
FT 12-DEC-1996.
FT
FT 04-JUN-1996; 96WO-EP002432.
FT
FT 06-JUN-1995; 95US-00467397.
FT
FT (HOFF ) HOFFMANN LA ROCHE & CO AG F.
FT (HYBR-) HYBRIDON INC.
FT
FT Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
FT Roberts NA, Roberts PC, Slade A;
FT WPI; 1997-043124/04.
FT
FT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
FT used in the detection and treatment of HBV infection.
FT
FT Claim 5; Page 15; 81pp; English.
FT
FT The present sequence represents a synthetic oligonucleotide HBV-89Mb
FT which contains a sequence which is complementary to at least two non-
FT contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
FT antisense oligonucleotide may be used to detect the presence of HBV in a
FT sample. The antisense oligonucleotide, and oligonucleotides complementary
FT to a portion of the HBV RNA, may be used for inhibiting HBV replication
FT
CC in a cell or for the treatment of HBV infection
XX
SQ Sequence 30 BP; 12 A; 3 C; 9 G; 2 T; 4 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAAAGAGAGAUU 20
Db 1 GACAUGAACAAAGAGAGAUU 20
RESULT 7
ADC64742/c
ID ADC64742 standard; RNA; 39 BP.
XX
AC ADC64742;
XX
DT 18-DEC-2003 (first entry)
XX
DE Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX
KW screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX
OS Synthetic.
OS Hepatitis B virus.
XX
PN KR2002007891-A.
XX
PD 29-JAN-2002.
XX
PF 19-JUL-2000; 2000KR-00041420.
XX
PR 19-JUL-2000; 2000KR-00041420.
XX
PA (MOGA-) MOGAM BIOTECHNOLOGY INST.
PA (VIRO-) VIROGEN CO LTD.
XX
PI Ji HJ, Jung SI, Kim YC, Min MG, Ryu WS, Yoon GS;
XX WPI; 2003-309015/30.
XX
PT Screening of antiviral agents by protein-priming activity of hepatitis B
PT virus DNA polymerase.
XX
PS Disclosure; Page 12; 13pp; Korean.
XX
CC The present invention describes a method of screening for an antiviral
CC agent by the protein-priming activity of hepatitis B virus (HBV) DNA
CC polymerase. Also described is developing an antiviral agent with a high
CC selectivity to HBV which can be used for high-throughput screening. The
CC present sequence represents an RNA oligonucleotide which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 39 BP; 5 A; 13 C; 3 G; 0 T; 18 U; 0 Other;
Query Match 100.0%; Score 20; DB 10; Length 39;
Best Local Similarity 80.0%; Pred. No. 10;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAUGAACAAAGAGAGAUU 20
Db 37 GACATGACAAAGAGATGATT 18
RESULT 8
AAD09094/c
ID AAD09094 standard; DNA; 87 BP.
XX
AC AAD09094;
XX
DT 04-SEP-2001 (first entry)
```

```

XX DE Hepatitis B virus FRI strain genotype G HBeAg DNA fragment.
XX DE
XX KW HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
XX KW preS2; surface antigen; HBSAg; HBx protein; vaccine; HBeAg;
XX KW liver disease; hepatitis; liver cancer; HBCAg; core antigen; ds.
XX OS
XX OS Hepatitis B virus.
XX PN WO200138498-A2.
XX XX
XX FD 31-MAY-2001.
XX XX
XX PF 21-NOV-2000; 2000WO-US032108.
XX PF
XX PR 24-NOV-1999; 99US-0167206P.
XX PR
XX PA (PHAR-) PHARMASSET INC.
XX PA (INNO-) INNOGENETICS NV.
XX XX
XX PI Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M;
XX PI Rossau R;
XX DR
XX DR MPI; 2001-367676/38.
XX XX
XX PT Novel hepatitis B virus genotype G, nucleic acids encoding virus,
XX PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
XX PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX PS Claim 6; Page 57; 84pp; English.
XX CC The present invention relates to hepatitis B virus (HBV) strain FRI,
XX CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
XX CC PreS2 and surface antigen HBSAg) and HBx proteins. HBV genotype G nucleic
XX CC acids and polypeptides are useful for diagnosing, prognosing and treating
XX CC infections caused by HBV genotype G. They can be used in a vaccine to
XX CC treat or prevent HBV genotype G infection. The HBV genotype G derived
XX CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
XX CC sample or diagnosis of HBV genotype G infection. The presence of HBV
XX CC genotype G statistically correlates with the presence of liver damage
XX CC and/or liver cancer in the subject. The HBV genotype G core insert
XX CC peptide encoding nucleic acid is useful for designing monitoring assays
XX CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
XX CC and HBeAg (genotype G e antigen) in patients infected with HBV. The
XX CC antibodies or antigens of HBV genotype G are useful for identifying a
XX CC stage of liver disease caused by HBV genotype G. The present sequence is
XX CC a hepatitis B virus (HBV) strain FRI, genotype G DNA fragment encoding e
XX CC antigen (HBeAg)
XX SQ Sequence 87 BP; 14 A; 24 C; 17 G; 32 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 87;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGUAAU 20
DB 43 GACATGAACAGAGATGATT 24

RESULT 9
AAD09093/c
ID AAD09093 standard; DNA; 129 BP.
XX AC
XX AC AAD09093;
XX DT
XX DT 04-SEP-2001 (first entry)
XX DE Hepatitis B virus FRI strain genotype G DNA fragment #1.
XX KW HBV genotype G; precore; HBpol; polymerase; envelope protein; preS1;
XX KW preS2; surface antigen; HBSAg; HBx protein; vaccine; liver disease;
XX KW hepatitis; liver cancer; HBCAg; core antigen; ds.

```

```

XX OS Hepatitis B virus.
XX PN WO200138498-A2.
XX XX
XX PD 31-MAY-2001.
XX XX
XX PF 21-NOV-2000; 2000WO-US032108.
XX PF
XX PR 24-NOV-1999; 99US-0167206P.
XX PR
XX PA (PHAR-) PHARMASSET INC.
XX PA (INNO-) INNOGENETICS NV.
XX XX
XX PI Stuyver L, Schinazi R, De Gendt S, Van Geyt C, Zoulim F, Fried M;
XX PI Rossau R;
XX DR
XX DR MPI; 2001-367676/38.
XX XX
XX PT Novel hepatitis B virus genotype G, nucleic acids encoding virus,
XX PT polypeptides encoded by nucleic acids, useful for preparing vaccine to
XX PT treat or prevent the hepatitis B virus genotype G infection in a subject.
XX PS Claim 5; Page 57; 84pp; English.
XX CC The present invention relates to hepatitis B virus (HBV) strain FRI,
XX CC genotype G DNA encoding PreCore/Core protein, HBpol, envelope (PreS1,
XX CC PreS2 and surface antigen HBSAg) and HBx proteins. HBV genotype G nucleic
XX CC acids and polypeptides are useful for diagnosing, prognosing and treating
XX CC infections caused by HBV genotype G. They can be used in a vaccine to
XX CC treat or prevent HBV genotype G infection. The HBV genotype G derived
XX CC nucleic acids and antibodies are useful for detecting HBV genotype G in a
XX CC sample or diagnosis of HBV genotype G infection. The presence of HBV
XX CC genotype G statistically correlates with the presence of liver damage
XX CC and/or liver cancer in the subject. The HBV genotype G core insert
XX CC peptide encoding nucleic acid is useful for designing monitoring assays
XX CC to study and predict the evolution of anti-HBe and anti-HBc antibodies
XX CC and HBeAg (genotype G e antigen) in patients infected with HBV. The
XX CC antibodies or antigens of HBV genotype G are useful for identifying a
XX CC stage of liver disease caused by HBV genotype G. The present sequence is
XX CC a hepatitis B virus (HBV) strain FRI, genotype G DNA fragment
XX SQ Sequence 129 BP; 25 A; 32 C; 26 G; 46 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 4; Length 129;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGAGUAAU 20
DB 43 GACATGAACAGAGATGATT 24

RESULT 10
AAD27422/c
ID AAD27422 standard; DNA; 639 BP.
XX AC
XX AC AAD27422;
XX DT
XX DT 18-APR-2002 (first entry)
XX DE Hepatitis B virus (HBV) core antigen (HBCAg) encoding DNA #1.
XX KW Hepatitis B virus; HBV; core antigen; HBCAg; immune system; typhoid;
XX KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
XX KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
XX KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
XX KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
XX KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
XX KW antiprotozoal; ds.
XX OS Hepatitis B virus.
XX OS

```

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-61

Perfect score: 20

Sequence: 1 gacgaagaacagaagauu 20

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	92.0	763	9	CG381974
C 2	18.4	92.0	773	9	CG381984
C 3	18.4	92.0	846	9	CG373212
C 4	18.4	92.0	892	9	CG373225
C 5	17.4	87.0	280	9	EX288914
C 6	17.4	87.0	305	6	CA520799
C 7	17.4	87.0	313	7	CF906991
C 8	17.4	87.0	423	5	EX837806
C 9	17.4	87.0	426	8	AQ183243
C 10	17.4	87.0	442	7	CN958899
C 11	17.4	87.0	450	8	AZ654527
C 12	17.4	87.0	509	3	AK015232
C 13	17.4	87.0	527	8	AQ961278
C 14	17.4	87.0	570	7	CF198522
C 15	17.4	87.0	574	7	CF909462
C 16	17.4	87.0	616	7	CK517114
C 17	17.4	87.0	628	8	BH366000
C 18	17.4	87.0	631	2	BE388774
C 19	17.4	87.0	652	8	BZ898091
C 20	17.4	87.0	682	9	CE163565
C 21	17.4	87.0	683	7	CO817689
C 22	17.4	87.0	721	8	AQ961277
C 23	17.4	87.0	794	9	CL809904
C 24	17.4	87.0	821	2	BF678287

C	25	17.4	87.0	859	8	CC090167
C	26	17.4	87.0	870	8	CC131380
C	27	17.4	87.0	885	2	BF541940
C	28	17.4	87.0	903	3	CC068329
C	29	17.4	87.0	1340	3	CNS0A5V5
C	30	17.4	87.0	1472	2	AW760013
C	31	17.4	87.0	532	1	AL819446
C	32	17.4	87.0	577	4	BI378081
C	33	17.4	87.0	687	9	AG140746
C	34	17.4	87.0	730	5	EX114353
C	35	17.4	87.0	768	7	CC786410
C	36	17.4	87.0	778	7	CO368799
C	37	17.4	87.0	808	9	CL543431
C	38	17.4	87.0	1013	9	CNS06RAQ
C	39	16.8	84.0	124	4	BI127956
C	40	16.8	84.0	170	4	BI128183
C	41	16.8	84.0	195	4	BG125265
C	42	16.8	84.0	195	4	BG733602
C	43	16.8	84.0	218	2	BE428564
C	44	16.8	84.0	234	4	BI473621
C	45	16.8	84.0	251	8	BZ385056

ALIGNMENTS

RESULT 1	CG381974	763 bp	DNA	linear	GSS 26-AUG-2003
LOCUS	CG381974	CG1BK24TH ZM 0.7_1.5_KB	Zea mays	genomic clone ZMMBMA0724C23,	
DEFINITION	CG381974	genomic survey sequence.			
ACCESSION	CG381974	GI:34299241			
VERSION	CG381974.1	GSS			
KEYWORDS	CG381974.1	Zea mays			
SOURCE	CG381974.1	Zea mays			
ORGANISM	CG381974.1	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.			
REFERENCE	CG381974.1	1 (bases 1 to 763)			
AUTHORS	CG381974.1	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.			
TITLE	CG381974.1	Consortium for Maize Genomics			
JOURNAL	CG381974.1	Unpublished (2002)			
COMMENT	CG381974.1	Other GSSs: CG1BK24TV Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends.			

FEATURES

source	CG381974	Location/Qualifiers
/organism="Zea mays"		
/mol_type="genomic DNA"		
/strain="B73"		
/db_xref="taxon:4577"		
/clone="ZMMBMA0724C23"		
/note="Vector: pBCKS-1; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"		

ORIGIN

Query Match	92.0%	Score 18.4;	DB 9;	Length 763;
Best Local Similarity	75.0%	Pred. No. 4.2e+02;		
Matches	15;	Conservative	4;	Mismatches
			1;	Indels
			0;	Gaps
			0;	
Qy	1	GACUACACAGAGAUU	20	
Db	572	GCATGACACAGATGATT	591	

```

RESULT 2
CG381984/c
LOCUS       OGLBK24TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBma0724C23,
            773 bp    DNA    linear    GSS 26-AUG-2003
DEFINITION  genomic survey sequence.
ACCESSION   CG381984
VERSION     CG381984.1  GI:34299251
KEYWORDS    GSS.
SOURCE      Zea mays
            Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 773)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Other GSSs: CG1BK24TH
            Contact: Cathy Whitelaw
            TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..773
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
Query Match      92.0%; Score 18.4; DB 9; Length 773;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

CY 1 GACAUACAACAGAGAUUU 20
    |||:|||||:|||||:|||||:
Db 668 GGCATGACACAGAGATGATT 649

RESULT 3
CG373212/c
LOCUS       CG1CZ28TH_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBma0734E08,
            846 bp    DNA    linear    GSS 26-AUG-2003
DEFINITION  genomic survey sequence.
ACCESSION   CG373212
VERSION     CG373212.1  GI:34290479
KEYWORDS    GSS.
SOURCE      Zea mays
            Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 846)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Other GSSs: CG1CZ28TV
            Contact: Cathy Whitelaw
            TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..846
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
Query Match      92.0%; Score 18.4; DB 9; Length 846;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

CY 1 GACAUACAACAGAGAUUU 20
    |||:|||||:|||||:|||||:
Db 668 GGCATGACACAGAGATGATT 649

RESULT 4
CG373225
LOCUS       CG1CZ28TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBma0734E08,
            892 bp    DNA    linear    GSS 26-AUG-2003
DEFINITION  genomic survey sequence.
ACCESSION   CG373225
VERSION     CG373225.1  GI:34290492
KEYWORDS    GSS.
SOURCE      Zea mays
            Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 892)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Other GSSs: CG1CZ28TH
            Contact: Cathy Whitelaw
            TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..892
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
Query Match      92.0%; Score 18.4; DB 9; Length 892;
Best Local Similarity 75.0%; Pred. No. 4.4e+02;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

CY 1 GACAUACAACAGAGAUUU 20
    |||:|||||:|||||:|||||:

```

```

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..846
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="ZMMBma0734E08"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
Query Match      92.0%; Score 18.4; DB 9; Length 846;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

CY 1 GACAUACAACAGAGAUUU 20
    |||:|||||:|||||:|||||:
Db 540 GGCATGACACAGAGATGATT 521

RESULT 4
CG373225
LOCUS       CG1CZ28TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBma0734E08,
            892 bp    DNA    linear    GSS 26-AUG-2003
DEFINITION  genomic survey sequence.
ACCESSION   CG373225
VERSION     CG373225.1  GI:34290492
KEYWORDS    GSS.
SOURCE      Zea mays
            Zea mays
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 892)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
            Other GSSs: CG1CZ28TH
            Contact: Cathy Whitelaw
            TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..892
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="ZMMBma0734E08"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
Query Match      92.0%; Score 18.4; DB 9; Length 892;
Best Local Similarity 75.0%; Pred. No. 4.4e+02;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

CY 1 GACAUACAACAGAGAUUU 20
    |||:|||||:|||||:|||||:

```

Query Match	87.0%	Score 17.4;	DB 9;	Length 280;
Best Local Similarity	78.9%;	Pred. No. 1.1e-03;		
Matches 15;	Conservative	3;	Mismatches 1;	Indels 0;
	Gaps	0;		
MEDLINE				
	21423036	PUBMED		
	11544199			
COMMENT				
Contact: Dawood B. Dudekula				
Laboratory of Genetics				

AZ654527	450 bp	DNA	linear	GSS 14-DEC-2000	
LOCUS					
DEFINITION					
cloned UUC1M library Mus musculus genomic					
clone UUC1M0528D16 R, genomic sequence.					
AK015232	509 bp	mRNA	linear	AK015232	
LOCUS					
DEFINITION					
Mus musculus adult male testis cDNA, RIKEN full-length enriched					
library, clone:4930429C20 product:unclassified, full insert					
AK015232	509 bp	mRNA	linear	AK015232	
LOCUS					
DEFINITION					
Mus musculus adult male testis cDNA, RIKEN full-length enriched					
library, clone:4930429C20 product:unclassified, full insert					

COMMENT

sheared to 0.9-1 Kbp before ligation."

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 527;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACAUGAACAGAGGAUGAUU 20
|||||
Db 160 ACATGAACAAGAGGATT 142

RESULT 14

CF198522 570 bp mRNA linear EST 01-AUG-2003
LOCUS EST0117 Tamarix androssowii leaf Tamarix androssowii cDNA, mRNA
DEFINITION sequence.

ACCESSION CF198522
VERSION CF198522.1 GI:33392895
KEYWORDS EST.
SOURCE Tamarix androssowii

ORGANISM Tamarix androssowii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Tamaricaceae; Tamarix.

REFERENCE

AUTHORS Wang, Y., Yang, C., Jiang, J., Liu, G., Wu, J. and Liu, Z.
TITLE EST acquired from cDNA library of Tamarix androssowii treated with NaHCO3

JOURNAL

COMMENT Unpublished (2003)

Contact: Yucheng Wang
Forestry Source and Environment College
Northeast Forestry University
Hexing 26, Harbin, Heilongjiang, 150040, P.R. China
Tel: 086-451-2190607
Email: WANGYUCHENG1029@YAHOO.COM.CN.

FEATURES

source 1..570
/organism="Tamarix androssowii"
/mol_type="mRNA"
/db_xref="taxon:189785"
/tissue_type="leaf"
/clone_lib="Tamarix androssowii leaf"

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 570;
Best Local Similarity 73.7%; Pred. No. 1.2e+03;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACAUGAACAGAGGAUGAUU 20
|||||
Db 143 ACATGAACAAGAGGTGATT 161

RESULT 15

CF909462/c A0536H03-5 NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)
LOCUS A0536H03-5 NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)
DEFINITION Mus musculus cDNA clone NIA:A0536H03 IMAGE:30746294 5', mRNA
sequence.

ACCESSION CF909462
VERSION CF909462.1 GI:38180399
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 574)

REFERENCE

AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

JOURNAL Genome Res. 11 (9), 1553-1558 (2001)

MEDLINE

PUBMED 21429098

COMMENT 11544199

Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdns@leuun.grc.nia.nih.gov
Plate: A0536 row: H column: 03
Seq primer: M13 Reverse
High quality sequence stop: 574
POLYA=No.

FEATURES

source

Location/Qualifiers 1..574

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H/He mice"
/db_xref="niaEST:A0536H03-5"
/db_xref="taxon:10090"
/clone="NIA:A0536H03 IMAGE:30746294"
/dev_stage="9-15C cells"
/lab_host="DH10B"
/clone_lib="NIA Mouse Mesenchymal Stem Cell cDNA Library (Long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: Salt; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Akihiro Umezawa (Keio University School of Medicine, Japan). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen]:
5'-pGACTAGTTCATGATCGAGCGCGCCCTTTTCTTTT-3' from 2.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.5 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 574;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACAUGAACAGAGGAUGAU 19

Db 510 GAAATGAACAAGAGATGAT 492

Search completed: March 17, 2005, 11:07:46
Job time : 1386.27 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-62

Perfect score: 20
Sequence: 1 taagggtcgauccauccg 20

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6	AR027817
2	20	100.0	30	6	AR027842
3	20	100.0	30	6	AR027843
4	20	100.0	81	6	I92348
5	20	100.0	174	14	S77749
6	20	100.0	253	14	AY329529
7	20	100.0	253	14	AY329562
8	20	100.0	253	14	AY329568
9	20	100.0	253	14	AY329573
10	20	100.0	253	14	AY329581
11	20	100.0	254	14	AF390000
12	20	100.0	333	14	HP8HBD
13	20	100.0	398	14	AB167603
14	20	100.0	398	14	AB167637
15	20	100.0	406	14	AB163815
16	20	100.0	406	14	AB163817
17	20	100.0	439	14	AY254503
18	20	100.0	488	14	AY274419
19	20	100.0	488	14	AY274420

C 20	20	100.0	488	14	AY274422	488	14	AY274422	Hepatitis
C 21	20	100.0	488	14	AY274427	488	14	AY274427	Hepatitis
C 22	20	100.0	488	14	AY274428	488	14	AY274428	Hepatitis
C 23	20	100.0	488	14	AY274429	488	14	AY274429	Hepatitis
C 24	20	100.0	488	14	AY274430	488	14	AY274430	Hepatitis
C 25	20	100.0	488	14	AY274431	488	14	AY274431	Hepatitis
C 26	20	100.0	488	14	AY274432	488	14	AY274432	Hepatitis
C 27	20	100.0	488	14	AY274433	488	14	AY274433	Hepatitis
C 28	20	100.0	488	14	AY274434	488	14	AY274434	Hepatitis
C 29	20	100.0	488	14	AY274436	488	14	AY274436	Hepatitis
C 30	20	100.0	493	14	S79556	493	14	S79556	X, prec (he
C 31	20	100.0	548	14	AY382500	548	14	AY382500	Hepatitis
C 32	20	100.0	548	14	AY382501	548	14	AY382501	Hepatitis
C 33	20	100.0	548	14	AY382502	548	14	AY382502	Hepatitis
C 34	20	100.0	548	14	AY382521	548	14	AY382521	Hepatitis
C 35	20	100.0	548	14	AY382522	548	14	AY382522	Hepatitis
C 36	20	100.0	548	14	AY382523	548	14	AY382523	Hepatitis
C 37	20	100.0	548	14	AY382524	548	14	AY382524	Hepatitis
C 38	20	100.0	548	14	AY382525	548	14	AY382525	Hepatitis
C 39	20	100.0	548	14	AY382526	548	14	AY382526	Hepatitis
C 40	20	100.0	548	14	AY382527	548	14	AY382527	Hepatitis
C 41	20	100.0	609	14	AF289954	609	14	AF289954	Hepatitis
C 42	20	100.0	609	14	AF289965	609	14	AF289965	Hepatitis
C 43	20	100.0	626	14	AY254500	626	14	AY254500	Hepatitis
C 44	20	100.0	639	6	AX278066	639	6	AX278066	Sequence
C 45	20	100.0	639	6	AX342485	639	6	AX342485	Sequence

ALIGNMENTS

RESULT 1
AR027817
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Sequence 15 from patent US 5856459.
AR027817
AR027817.1 GI:5938637
Unknown.
Unknown.

REFERENCE
AUTHORS
Mills, J.S.

1 (bases 1 to 20)
Frank, B.L., Roberts, P.C., Goodchild, J., Craig, J. Charles. and

OLIGONUCLEOTIDES SPECIFIC FOR HEPATITIS B VIRUS
Patent: US 5856459-A 15 05-JAN-1999;

LOCATION/QUALIFIERS
1..20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 85.0%; Pred. No. 14;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCCAUCC 20
|||||||:|:|:|:|:|:|

Db 1 TAAGGTCGATGTCATGCC 20
|||||||:|:|:|:|:|:|

RESULT 2

AR027842
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Sequence 40 from patent US 5856459.
AR027842
AR027842.1 GI:5938662
Unknown.
Unknown.

REFERENCE
AUTHORS
Frank, B.L., Roberts, P.C., Goodchild, J., Craig, J. Charles. and

1 (bases 1 to 30)
Frank, B.L., Roberts, P.C., Goodchild, J., Craig, J. Charles. and

OLIGONUCLEOTIDES SPECIFIC FOR HEPATITIS B VIRUS
Patent: US 5856459-A 15 05-JAN-1999;

LOCATION/QUALIFIERS
1..20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 85.0%; Pred. No. 14;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCCAUCC 20
|||||||:|:|:|:|:|:|

Db 1 TAAGGTCGATGTCATGCC 20
|||||||:|:~|:|:|:|:|:|

Qy

```

/db_xref="taxon:10407"
<1..158
/product="X protein"
/protein_id="AAQ95861.1"
/db_xref="GI:37625316"
/translation="STTDLEAYFKDVFKDWEELGEBEIRLMIFVLGGCRHKLVCAPAS
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95862.1"
/db_xref="GI:37625317"
/translation="MQLFHLCLIVISCTCTPTQASKLCLGLW"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCCAUGCC 20
|||||:|:|:|:|
Db 237 TAAGGTCGATGTCATGCC 218

RESULT 7
AV329562/c
LOCUS
DEFINITION
Hepatitis B virus isolate D273984E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329562
VERSION
AY329562.1 GI:37625413
KEYWORDS
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
PUBMED
15184419
AUTHORS
Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratorio
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
FEATURES
Location/Qualifiers
source
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolates="D273984E"
/db_xref="taxon:10407"
<1..158
/product="X protein"
/protein_id="AAQ95938.1"
/db_xref="GI:37625432"
/translation="STTDLEAYFKDCLFKDWBELGELRLMIFVLGGCRHKLVCAPAP
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95939.1"
/db_xref="GI:37625433"
/translation="MQLFHLCLIIISCSCTPTQASKLCLGLW"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCCAUGCC 20
|||||:|:|:|:|
Db 237 TAAGGTCGATGTCATGCC 218

RESULT 9
AV329573/c
LOCUS
DEFINITION
Hepatitis B virus isolate D604917E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329573
VERSION
AY329573.1 GI:37625446
KEYWORDS
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

```

```

/isolate="D639472E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95964.1"
/db_xref="GI:37625471"
/translation="STTDLEAYFKDCLFKDWELGEBELRLIFVLGGCRHKLVCAPAP
CNFFTSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95965.1"
/db_xref="GI:37625472"
/translation="WQLFHLCLIISCSCTVQASKLCLGWL"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. NO. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCC 20
|||||:|:|:|
Db 237 TAAGGTCGATGCC 218
|||||:|:|:|

RESULT 11
AF390000/c
LOCUS
DEFINITION
294 bp DNA linear VRL 06-MAR-2002
Hepatitis B virus isolate D3 X protein gene, partial cds; and
nonfunctional precore/core protein gene, partial sequence.
ACCESSION
AF390000
VERSION
AF390000.1 GI:16266099
KEYWORDS
Hepatitis B virus
SOURCE
Hepatitis B virus
ORGANISM
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 294)
Castro,L.D., Niel,C. and Gomes,S.A.
Low frequency of mutations in the core promoter and precore regions
of Hepatitis B virus in anti-HBe positive Brazilian carriers
BMC Microbiol. 1 (1), 10 (2001)
11472634
2 (bases 1 to 294)
De Castro,L., Niel,C. and Gomes,S.A.
Direct Submission
AUTHORS
Submitted (11-JUN-2001) Virology, FIOCRUZ, Av. Brasil 4365, Rio de
Janeiro, RJ 21045-900, Brazil
JOURNAL
Location/Qualifiers
FEATURES
source
1..294
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D3"
/db_xref="taxon:10407"
<1..119
/codon_start=3
/product="X protein"
/protein_id="AAL16752.1"
/db_xref="GI:16266100"
/translation="FKDWELGDSRLMIYVLGGCRHKLVCAPAPCNFFTSA"
95..>294
/misc_feature
/note="nonfunctional precore/core protein due to mutation"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 294;
Best Local Similarity 85.0%; Pred. NO. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCC 20
|||||:|:|:|
Db 198 TAAGGTCGATGCC 179
|||||:|:|:|

```

```
RESULT 12
HPBBED
LOCUS          333 bp      DNA      linear      VRL 28-JUN-1996
DEFINITION    Hepatitis B virus precore and core protein gene, 5' end of cds.
ACCESSION     L12359
VERSION       L12359.1 GI:306267
KEYWORDS      HBcAg protein; HBcAg protein; core protein; nucleotide binding
SOURCE        Hepatitis B virus
ORGANISM      Hepatitis B virus
REFERENCE     1 (sites)
AUTHORS      Tong,S.P., Li,J.S., Vitvitski,L. and Trepo,C.
TITLE        Active hepatitis B virus replication in the presence of anti-HBe is
JOURNAL      associated with viral variants containing an inactive pre-C region
MEDLINE      Virology 176 (2), 596-603 (1990)
PUBMED       90266476
REFERENCE     2 (bases 1 to 333)
AUTHORS      Li,J.S., Tong,S.P., Wen,Y.M., Vitvitski,L., Zhang,Q. and Trepo,C.
TITLE        Hepatitis B virus genotype A rarely circulates as an HBe-minus
JOURNAL      mutant: possible contribution of a single nucleotide in the precore
MEDLINE      region
PUBMED       2345966
REFERENCE     3 (bases 1 to 333)
AUTHORS      J. Virol. 67 (9), 5402-5410 (1993)
TITLE        Hepatitis B virus DNA.
JOURNAL      Original source text: Hepatitis B virus DNA.
MEDLINE      Location/Qualifiers
PUBMED       8350403
COMMENT       Original source text: Hepatitis B virus DNA.
FEATURES      Location/Qualifiers
source        1..333
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10407"
              /note="genotype D; from French HBeAg- patient"
              1..>333
              /standard_name="HBeAg"
              /pseudo
              /codon_start=1
              /product="precore protein"
              82..84
              /note="HBe-abolishing"
              88..>333
              /standard_name="HBCAb"
              /codon_start=1
              /db_xref="GI:306268"
              /db_xref="GI:306268"
              /translation="MDIDPYKEFGATVELTFLPDPFPSVRDLDTAAALVREALS
              PECHSHHTALRQALICWGLMTLATWVGANLDDPASR"
              CDS
              variation
              82..84
              /product="precore protein"
              88..>333
              /standard_name="HBCAb"
              /codon_start=1
              /db_xref="GI:306268"
              /db_xref="GI:306268"
              /translation="MDIDPYKEFGATVELTFLPDPFPSVRDLDTAAALVREALS
              PECHSHHTALRQALICWGLMTLATWVGANLDDPASR"
              ORIGIN
              Query Match          100.0%; Score 20; DB 14; Length 333;
              Best Local Similarity 85.0%; Pred. No. 17;
              Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
              QY 1 TAAGGTCGAUGUCCAUGCC 20
              |||||
              Db 104 TAAGGTCGATGTCATGCC 85
              |||||
              RESULT 13
              AB167603/c
              LOCUS          398 bp      DNA      linear      VRL 01-OCT-2004
              DEFINITION    Hepatitis B virus gene for polyprotein, partial cds, clone: NEP75.
              ACCESSION     AB167603
              VERSION       AB167603.1 GI:53148166
              KEYWORDS      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
              SOURCE        Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
              ORGANISM      A Case-control Study for Differences among Hepatitis B Virus
              Infections of Genotypes A (Subtypes Aa and Ae) and D
              Unpublished
              Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
              Submitted (15-MAR-2004) Yasuhiro Tanaka, Nagoya City University
              Graduate School of Medical Sciences, Department of Clinical
              Molecular Informative Medicine, 1 Kawasumi, Mizuho-cho, Mizuho-ku,
              Nagoya, Aichi 467-8601, Japan (E-mail:ytanaka@med.nagoya-cu.ac.jp,
              Tel:81-52-853-8292, Fax:81-52-842-0021)
              Tel:81-52-853-8292, Fax:81-52-842-0021)
              Location/Qualifiers
              source        1..398
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10407"
              /clone="NEP75"
              184..>398
              /note="precore-core region"
              /codon_start=1
              /product="polyprotein"
              /protein_id="BAD52175.1"
              /db_xref="GI:53148167"
              /translation="MQLFHLCLIIISCTPTFOASKLCLGLWGMMDIDPYKEFGATVEL
              LSFPLSDPPFPSVRDLDTASALYREAL"
              CDS
              184..>398
              /note="precore-core region"
              /codon_start=1
              /product="polyprotein"
              /protein_id="BAD52209.1"
              /db_xref="GI:53148361"
              /translation="MQLFHLCLIIISCTPTFOASKLCLGLWGMMDIDPYKEFGATVEL
              LSFPLSDPPFPSVRDLDTASALYREAL"
              ORIGIN
              Query Match          100.0%; Score 20; DB 14; Length 398;
              Best Local Similarity 85.0%; Pred. No. 18;
              Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
              QY 1 TAAGGTCGAUGUCCAUGCC 20
              |||||
              Db 287 TAAGGTCGATGTCATGCC 268
              |||||
              RESULT 14
              AB167637/c
              LOCUS          398 bp      DNA      linear      VRL 01-OCT-2004
              DEFINITION    Hepatitis B virus gene for polyprotein, partial cds, clone: NEP27.
              ACCESSION     AB167637
              VERSION       AB167637.1 GI:53148360
              KEYWORDS      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
              SOURCE        Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
              ORGANISM      A Case-control Study for Differences among Hepatitis B Virus
              Infections of Genotypes A (Subtypes Aa and Ae) and D
              Unpublished
              Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
              Submitted (15-MAR-2004) Yasuhiro Tanaka, Nagoya City University
              Graduate School of Medical Sciences, Department of Clinical
              Molecular Informative Medicine, 1 Kawasumi, Mizuho-cho, Mizuho-ku,
              Nagoya, Aichi 467-8601, Japan (E-mail:ytanaka@med.nagoya-cu.ac.jp,
              Tel:81-52-853-8292, Fax:81-52-842-0021)
              Tel:81-52-853-8292, Fax:81-52-842-0021)
              Location/Qualifiers
              source        1..398
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10407"
              /clone="NEP27"
              184..>398
              /note="precore-core region"
              /codon_start=1
              /product="polyprotein"
              /protein_id="BAD52209.1"
              /db_xref="GI:53148361"
              /translation="MQLFHLCLIIISCTPTFOASKLCLGLWGMMDIDPYKEFGATVEL
              LSFPLSDPPFPSVRDLDTASALYREAL"
              CDS
              184..>398
              /note="precore-core region"
              /codon_start=1
              /product="polyprotein"
              /protein_id="BAD52209.1"
              /db_xref="GI:53148361"
              /translation="MQLFHLCLIIISCTPTFOASKLCLGLWGMMDIDPYKEFGATVEL
              LSFPLSDPPFPSVRDLDTASALYREAL"
              ORIGIN
              Query Match          100.0%; Score 20; DB 14; Length 398;
              Best Local Similarity 85.0%; Pred. No. 18;
              Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
              QY 1 TAAGGTCGAUGUCCAUGCC 20
              |||||
              Db 287 TAAGGTCGATGTCATGCC 268
              |||||
              RESULT 15
              AB167603/c
              LOCUS          398 bp      DNA      linear      VRL 01-OCT-2004
              DEFINITION    Hepatitis B virus gene for polyprotein, partial cds, clone: NEP75.
              ACCESSION     AB167603
              VERSION       AB167603.1 GI:53148166
              KEYWORDS      Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
              SOURCE        Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
              ORGANISM      A Case-control Study for Differences among Hepatitis B Virus
              Infections of Genotypes A (Subtypes Aa and Ae) and D
              Unpublished
              Tanaka,Y., Hasegawa,I., Kato,T., Orito,E. and Mizokami,M.
              Submitted (15-MAR-2004) Yasuhiro Tanaka, Nagoya City University
              Graduate School of Medical Sciences, Department of Clinical
              Molecular Informative Medicine, 1 Kawasumi, Mizuho-cho, Mizuho-ku,
              Nagoya, Aichi 467-8601, Japan (E-mail:ytanaka@med.nagoya-cu.ac.jp,
              Tel:81-52-853-8292, Fax:81-52-842-0021)
              Tel:81-52-853-8292, Fax:81-52-842-0021)
              Location/Qualifiers
              source        1..398
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10407"
              /clone="NEP75"
              184..>398
              /note="precore-core region"
              /codon_start=1
              /product="polyprotein"
              /protein_id="BAD52175.1"
              /db_xref="GI:53148167"
              /translation="MQLFHLCLIIISCTPTFOASKLCLGLWGMMDIDPYKEFGATVEL
              LSFPLSDPPFPSVRDLDTASALYREAL"
              CDS
              184..>398
              /note="precore-core region"
              /codon_start=1
              /product="polyprotein"
              /protein_id="BAD52175.1"
              /db_xref="GI:53148167"
              /translation="MQLFHLCLIIISCTPTFOASKLCLGLWGMMDIDPYKEFGATVEL
              LSFPLSDPPFPSVRDLDTASALYREAL"
              ORIGIN
              Query Match          100.0%; Score 20; DB 14; Length 398;
              Best Local Similarity 85.0%; Pred. No. 18;
              Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
              QY 1 TAAGGTCGAUGUCCAUGCC 20
              |||||
              Db 287 TAAGGTCGATGTCATGCC 268
              |||||
```

Search completed: March 17, 2005, 08:14:17

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: us-08-901-612a-62

Perfect score: 20
Sequence: 1 taagggtcgauguccgcc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	2	AAT72572 Hepatitis
2	20	100.0	20	2	AAT72573 Hepatitis
3	20	100.0	30	2	AAT72619 Hepatitis
4	20	100.0	30	2	AAT72620 Hepatitis
5	20	100.0	30	2	AAT72618 Hepatitis
6	20	100.0	30	2	AAT72621 Hepatitis
C 7	20	100.0	31	10	ADG64743 Hepatitis
C 8	20	100.0	639	6	AAD27422 Hepatitis
C 9	20	100.0	639	6	AAD31509 Hepatitis
C 10	20	100.0	663	3	AAV71734 HBV fusio
C 11	20	100.0	669	12	ADO07220 Hepatitis
C 12	20	100.0	1334	2	AAV82691 Fulminant
C 13	20	100.0	1395	2	AAV82688 Fulminant
C 14	20	100.0	1400	2	AAV82687 Fulminant
C 15	20	100.0	1445	2	AAV82692 Fulminant
C 16	20	100.0	1445	2	AAV82685 Fulminant
C 17	20	100.0	1445	2	AAV82690 Fulminant
C 18	20	100.0	1445	2	AAV82684 Fulminant
C 19	20	100.0	1500	2	AAV82686 Fulminant
C 20	20	100.0	1500	2	AAV82689 Fulminant

C 21	20	100.0	2342	1	AAH93072 Sequence
C 22	20	100.0	3182	6	AAD31765 Hepatitis
C 23	20	100.0	3182	9	ACA62422 Hepatitis
C 24	20	100.0	3182	10	AAD60866 Hepatitis
C 25	20	100.0	5618	2	AAQ88310 Plasmid p
C 26	20	100.0	7991	6	AAH16094 HBV viral
C 27	20	100.0	8007	6	AAH16092 HBV viral
C 28	20	100.0	8717	6	AAH16093 HBV viral
C 29	19	95.0	34	10	ADJ94544 SDMCORE d
C 30	19	95.0	34	10	ADJ94545 SDMCORE d
C 31	19	95.0	39	13	ADR89273 Lab-on-ch
C 32	19	95.0	39	13	ADR89266 Lab-on-ch
C 33	18.8	94.0	30	2	AAQ45813 HBV ampli
C 34	18.8	94.0	30	2	AAV07810 HBV.D46 a
C 35	18.8	94.0	30	2	AAV83039 Amplifier
C 36	18.8	94.0	50	2	AAQ06723 :HBV.L1A2
C 37	18.4	92.0	22	10	ADG46961 PCR prime
C 38	18.4	92.0	22	11	ADM83206 PCR prime
C 39	18.4	92.0	24	6	ABK44212 B cell ep
C 40	18.4	92.0	24	6	ABK67439 Primer #1
C 41	18.4	92.0	24	6	ABK67506 Hepatitis
C 42	18.4	92.0	24	10	ADE80023 Primer fo
C 43	18.4	92.0	24	10	ADE10976 Chimeric
C 44	18.4	92.0	24	10	ADG46965 PCR prime
C 45	18.4	92.0	24	11	ADM83210 PCR prime

ALIGNMENTS

RESULT 1
AAT72572
ID AAT72572 standard; DNA; 20 BP.
XX
AC AAT72572;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV92b.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..20
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"

XX W09639502-A1.
XX
XX 12-DEC-1996.
XX
XX 04-JUN-1996; 96WO-EP002432.
XX
XX 06-JUN-1995; 95US-00467397.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX (HYBR-) HYBRIDON INC.
XX
XX Craig CU, Frank BL, Goodchild J, Jupp R, Kiluskie RE, Mills JS;
XX Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
XX used in the detection and treatment of HBV infection.
XX
XX Claim 1; Page 12; 81pp; English.

XX The present sequence represents a synthetic oligonucleotide HBV92b which
XX is complementary to a portion of the hepatitis B virus (HBV) RNA. The
XX antisense oligonucleotide may be used to detect the presence of HBV in a

CC sample. The antisense oligonucleotide, and oligonucleotides containing a
 CC sequence which is complementary to at least two non- contiguous regions
 CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
 CC cell or for the treatment of HBV infection

XX SQ Sequence 20 BP; 4 A; 5 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;
 Best Local Similarity 85.0%; Pred. No. 2.5;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGTCGAUGUCCAUCC 20
 |||||:|:|:|:
 Db 1 TAAGGGTCGATGCCATGCC 20

RESULT 2

AAAT72573
 ID AAAT72573 standard; DNA; 20 BP.

XX AC AAAT72573;

XX DT 03-SEP-1997 (first entry)

XX DE Hepatitis B virus RNA antisense oligonucleotide HBV92Mb.

XX HBV; HBV infection; inhibition; replication; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT misc_feature 1..20

FT /tag= a

FT /note= "Internucleotide linkages are phosphorothioate"

FT misc_RNA 11..20

FT /tag= b

FT /note= "2'-O-Me RNA"

FT modified_base 11

FT /tag= c

FT /mod_base= um

FT modified_base 12

FT /tag= d

FT /mod_base= gm

FT modified_base 13

FT /tag= e

FT /mod_base= um

FT modified_base 14

FT /tag= f

FT /mod_base= cm

FT modified_base 15

FT /tag= g

FT /mod_base= cm

FT modified_base 16

FT /tag= h

FT /mod_base= OTHER

FT /note= "2'-O-methyladenosine"

FT modified_base 17

FT /tag= i

FT /mod_base= um

FT modified_base 18

FT /tag= j

FT /mod_base= gm

FT modified_base 19

FT /tag= k

FT /mod_base= cm

FT modified_base 20

FT /tag= l

FT /mod_base= cm

XX WO9639502-A1.

XX PD 12-DEC-1996.

XX

PF 04-JUN-1996; 96WO-EP002432.

XX 06-JUN-1995; 95US-00467397.

XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX PA (HYBR-) HYBRIDON INC.

XX PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;

XX PI Roberts NA, Roberts PC, Slade A;

XX DR WPI; 1997-043124/04.

XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -

XX used in the detection and treatment of HBV infection.

XX Claim 1; Page 12; 81pp; English.

CC The present sequence represents a synthetic oligonucleotide HBV92Mb which
 CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
 CC antisense oligonucleotide may be used to detect the presence of HBV in a
 CC sample. The antisense oligonucleotide, and oligonucleotides containing a
 CC sequence which is complementary to at least two non- contiguous regions
 CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
 CC cell or for the treatment of HBV infection

XX SQ Sequence 20 BP; 4 A; 5 C; 6 G; 2 T; 3 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.5;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGTCGAUGUCCAUCC 20

|||||:|:|:|:
 Db 1 TAAGGGTCGAUGUCCAUCC 20

RESULT 3

AAAT72619

ID AAAT72619 standard; DNA; 30 BP.

XX AC AAAT72619;

XX DT 04-SEP-1997 (first entry)

XX DE Hepatitis B virus RNA antisense oligonucleotide HBV-90Mb.

XX HBV; HBV infection; inhibition; replication; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT misc_feature 1..30

FT /tag= a

FT /note= "Internucleotide linkages are phosphorothioate"

FT misc_RNA 1..10

FT /tag= b

FT /note= "2'-O-Me RNA"

FT modified_base 1

FT /tag= c

FT /mod_base= OTHER

FT modified_base 2

FT /tag= d

FT /mod_base= gm

FT modified_base 3

FT /tag= e

FT /mod_base= OTHER

FT modified_base 4

FT /tag= f

FT /mod_base= gm

FT modified_base 5

FT /tag= g

FT /mod_base= OTHER
 modified_base 6 /note= "2'-O-methyladenosine"
 FT /tag= h
 modified_base 7 /mod_base= um
 FT /tag= i
 modified_base 8 /mod_base= gm
 FT /tag= j
 modified_base 9 /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 modified_base 10 /tag= k
 FT /mod_base= um
 modified_base 11 /tag= l
 FT /mod_base= um
 misc_RNA 21..30
 FT /tag= m
 modified_base 21 /note= "2'-OMe RNA"
 FT /tag= n
 modified_base 22 /mod_base= um
 FT /tag= o
 modified_base 23 /mod_base= gm
 FT /tag= p
 modified_base 24 /mod_base= um
 FT /tag= q
 modified_base 25 /mod_base= cm
 FT /tag= r
 modified_base 26 /mod_base= cm
 FT /tag= s
 modified_base 27 /mod_base= OTHER
 FT /note= "2'-O-methyladenosine"
 modified_base 28 /tag= t
 FT /mod_base= um
 modified_base 29 /tag= u
 FT /mod_base= gm
 modified_base 30 /tag= v
 FT /mod_base= cm
 FT /tag= w
 modified_base /mod_base= cm
 XX
 PN WO9639502-A1.
 XX
 PD 12-DEC-1996.
 XX
 PP 04-JUN-1996; 96WO-EP002432.
 XX
 PR 06-JUN-1995; 95US-00467397.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 (HYBR-) HYBRIDON INC.
 XX
 PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
 Roberts NA, Roberts PC, Slade A;
 XX
 DR WPI; 1997-043124/04.
 XX
 PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 used in the detection and treatment of HBV infection.
 XX
 PS Claim 5; Page 15; 81pp; English.

XX The present sequence represents a synthetic oligonucleotide HBV-90Mb
 CC which contains a sequence which is complementary to at least two non-
 CC contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
 CC antisense oligonucleotide may be used to detect the presence of HBV in a
 CC sample. The antisense oligonucleotide, and oligonucleotides complementary
 CC to a portion of the HBV RNA, may be used for inhibiting HBV replication
 CC in a cell or for the treatment of HBV infection
 XX
 SQ Sequence 30 BP; 8 A; 5 C; 9 G; 2 T; 6 U; 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAAGGGTCGAUGUCCAUGCC 20
 Db 11 TAAGGGTCGAUGUCCAUGCC 30
 RESULT 4
 AAT72620
 ID AAT72620 standard; DNA; 30 BP.
 XX
 AC AAT72620;
 XX
 DT 04-SEP-1997 (first entry)
 XX
 DE Hepatitis B virus RNA antisense oligonucleotide HBV-91b.
 XX
 KW HBV; HBV infection; inhibition; replication; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 1..30
 FT /tag= a
 FT /note= "Internucleotide linkages are phosphorothioate"
 XX
 PN WO9639502-A1.
 XX
 PD 12-DEC-1996.
 XX
 PF 04-JUN-1996; 96WO-EP002432.
 XX
 PR 06-JUN-1995; 95US-00467397.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 (HYBR-) HYBRIDON INC.
 XX
 PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
 Roberts NA, Roberts PC, Slade A;
 XX
 DR WPI; 1997-043124/04.
 XX
 PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 used in the detection and treatment of HBV infection.
 XX
 PS Claim 5; Page 15; 81pp; English.
 XX
 CC The present sequence represents a synthetic oligonucleotide HBV-91b which
 CC contains a sequence which is complementary to at least two non-contiguous
 CC regions of a hepatitis B virus (HBV) nucleic acid. The antisense
 CC oligonucleotide may be used to detect the presence of HBV in a sample.
 CC The antisense oligonucleotide, and oligonucleotides complementary to a
 CC portion of the HBV RNA, may be used for inhibiting HBV replication in a
 CC cell or for the treatment of HBV infection
 XX
 SQ Sequence 30 BP; 7 A; 6 C; 11 G; 6 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 30;
 Best Local Similarity 85.0%; Pred. No. 2.6;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;


```

XX SQ Sequence 30 BP; 7 A; 6 C; 11 G; 3 T; 3 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAAGGTCGAUGUCCAUGCC 20
   |||||
Db 11 TAAAGGTCGAUGUCCAUGCC 30

RESULT 7
ADC64743/c
ID ADC64743 standard; RNA; 31 BP.
XX
AC ADC64743;
XX
DT 18-DEC-2003 (first entry)
XX
DE Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX
KW screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX
OS Synthetic.
OS Hepatitis B virus.
XX
PN KR2002007891-A.
XX
PD 29-JAN-2002.
XX
PF 19-JUL-2000; 2000KR-00041420.
XX
PR 19-JUL-2000; 2000KR-00041420.
XX
PA (MOGA-) MOGAM BIOTECHNOLOGY INST.
PA (VIRO-) VIROGEN CO LTD.
XX
PI Ji HJ, Jung SI, Kim YC, Min MG, Ryu WS, Yoon GS;
XX
WPI; 2003-309015/30.
XX
DR Screening of antiviral agents by protein-priming activity of hepatitis B
PT virus DNA polymerase.
XX
PS Disclosure; Page 12; 13pp; Korean.
XX
CC The present invention describes a method of screening for an antiviral
CC agent by the protein-priming activity of hepatitis B virus (HBV) DNA
CC polymerase. Also described is developing an antiviral agent with a high
CC selectivity to HBV which can be used for high-throughput screening. The
CC present sequence represents an RNA oligonucleotide which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 31 BP; 10 A; 6 C; 8 G; 0 T; 7 U; 0 Other;
Query Match 100.0%; Score 20; DB 10; Length 31;
Best Local Similarity 85.0%; Pred. No. 2.6;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAAGGTCGAUGUCCAUGCC 20
   |||||
Db 22 TAAAGGTCGATGTCATGCC 3

RESULT 8
AAD27422/c
ID AAD27422 standard; DNA; 639 BP.
XX
AC AAD27422;
XX
DT 18-APR-2002 (first entry)
XX

```

```

DE Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
XX
XX Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
KW antiprotozoal; ds.
XX
OS Hepatitis B virus.
XX
FH Key Location/Qualifiers
FT CDS 1..639
FT /tag= a
FT /product= "HBcAg"
XX
PN WO200198333-A2.
XX
PD 27-DEC-2001.
XX
PF 22-JUN-2001; 2001WO-GB002817.
XX
PR 22-JUN-2000; 2000GB-00015308.
PR 06-OCT-2000; 2000GB-00024544.
XX
PA (CELL-) CELLTech PHARM LTD.
XX
PI Page M, Li J, Pumpens P;
XX
WPI; 2002-098223/13.
DR P-PSDB; AAE17018.
XX
XX New proteins comprising a modified hepatitis B core antigen, useful as a
PT vaccine in prophylactic or therapeutic vaccination of the human or animal
PT body, particularly against hepatitis B virus infection.
XX
PS Disclosure; Page 38-39; 40pp; English.
XX
CC The invention relates to modified proteins comprising hepatitis B virus
CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine
CC repeats has been deleted and the protein comprising the C-terminal
CC cysteine of HBcAg. The deleted region may be replaced by an epitope from
CC a protein other than HBcAg, in which case the HBcAg acts as a carrier to
CC present the epitope to the immune system. This chimeric protein or its
CC nucleic acid is useful as a vaccine or in a method of prophylactic or
CC therapeutic vaccination of the human or animal body, particularly against
CC HBV. The nucleic acid encoding the protein may be used in gene therapy or
CC DNA vaccination protocols. The chimeric protein or its nucleic acid may
CC also be used as the basis of a prophylactic vaccine against a range of
CC diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
CC influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
CC immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
CC tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
CC diarrhoea, meningitis or gonorrhea. The present sequence is a DNA
CC encoding Hepatitis B virus core antigen (HBcAg)
XX
SQ Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 6; Length 639;
Best Local Similarity 85.0%; Pred. No. 3.3;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAAGGTCGAUGUCCAUGCC 20
   |||||
Db 104 TAAAGGTCGATGTCATGCC 85
   |||||

RESULT 9
AAD31509/c
ID AAD31509 standard; DNA; 639 BP.
XX
AC AAD31509;

```

```

XX DT 18-JUN-2002 (first entry)
XX DE Hepatitis B virus core antigen (HBcAg) encoding DNA.
XX KW Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
XX KW therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
XX KW polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX OS Hepatitis B virus.
XX PN WO200046376-A2.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 1..639
XX FT /*tag= a
XX FT /product= "Hbc protein"
XX FT sig_peptide 1..87
XX FT /*tag= b
XX FT mat_peptide 88..636
XX FT /*tag= c
XX FT /product= "Mature Hbc protein"
XX PN WO20017158-A1.
XX XX
XX PD 18-OCT-2001.
XX PF 09-APR-2001; 2001WO-GB001607.
XX XX
XX PR 07-APR-2000; 2000EP-00107118.
XX PA (MEDE-) MEDEVA EURO LTD.
XX PI Gehin A, Gilbert R, Stuart D, Rowlands D;
XX XX
XX DR WPI: 2002-239995/29.
XX DR P-PSDB; AAE19733.
XX XX
XX PT Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
XX PT prophylactic or therapeutic treatment of humans or animals against e.g.
XX PT HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
XX PT mouth disease.
XX PS Disclosure; Page 23-24; 27pp; English.
XX CC The present invention relates to hepatitis B virus (HBV) core antigen
XX CC (HBcAg) fusion proteins and polynucleotides encoding such proteins.
XX CC Sequences of the invention are useful in methods of prophylactic or
XX CC therapeutic vaccination or to manufacture medicaments for prophylactic or
XX CC therapeutic vaccination of the human or animal body against HBV, e.g.
XX CC against viral hepatitis. They are also useful as a prophylactic vaccine
XX CC against e.g. hepatitis c virus, influenza, polio, herpes, rabies,
XX CC acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
XX CC present sequence is a DNA encoding hepatitis B virus core antigen (HBcAg)
XX SQ Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 639;
Best Local Similarity 85.0%; Pred. No. 3.3;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGCCAUCCG 20
DB 104 TAAGGTCGATGTCATGCC 85

RESULT 10
AA71734/c
ID AAA71734 standard; cDNA; 663 BP.
XX AC AAA71734;
XX XX
XX DT 06-AUG-2003 (revised)
XX DT 08-JAN-2001 (first entry)
XX XX

```

```

DE DE HBV fusion protein comprising LHB and RGD encoding cDNA.
XX XX
XX KW Fusion protein; protein coat; virus-specific packaging signal; psi;
XX KW virus protein; cell permeability; cell-specific binding site; LHB;
XX KW large surface protein; core antigen; gene therapy; ss.
XX OS Hepatitis B virus.
XX OS Synthetic.
XX PN WO200046376-A2.
XX XX
XX PD 10-AUG-2000.
XX PF 04-FEB-2000; 2000WO-DE000363.
XX XX
XX PR 05-FEB-1999; 99DE-01004800.
XX XX
XX PA (HILD/) HILDT E.
XX XX
XX PI Hildt E, Hofschneider P;
XX XX
XX DR WPI: 2000-514959/46.
XX DR P-PSDB; AAB10596.
XX XX
XX PT Particle for cell-specific gene delivery, useful in gene therapy,
XX PT comprises nucleic acid in protein coat that includes a fusion protein of
XX PT viral protein, permeability peptide and cell-binding site.
XX PS Claim 16a; Fig 1; 34pp; German.
XX CC This invention describes a novel particle (A), comprising a protein coat
XX CC with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
XX CC including the sequence for a virus-specific packaging signal (psi) and a
XX CC structural gene. FP contains a virus protein (VP), a peptide (P) that
XX CC mediates cell permeability and a heterologous cell-specific binding site
XX CC (RGD). The invention also describes (1) producing (A) in which FP
XX CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
XX CC (RGD); (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
XX CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
XX CC containing the DNA of (d). The products of the invention are used in gene
XX CC therapy of cells and tissues, in vivo or ex vivo. This sequence encodes a
XX CC fusion protein which is described in the method of the invention.
XX CC (Updated on 06-AUG-2003 to correct OS field.)
XX SQ Sequence 663 BP; 154 A; 169 C; 152 G; 188 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 663;
Best Local Similarity 85.0%; Pred. No. 3.3;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGCCAUCCG 20
DB 107 TAAGGTCGATGTCATGCC 88

RESULT 11
AD007220/c
ID ADO07220 standard; DNA; 669 BP.
XX AC ADO07220;
XX XX
XX DT 15-JUL-2004 (first entry)
XX XX
XX DE Hepatitis B virus core antigen DNA.
XX KW HBcAg; immunomodulator; vaccine; gene; ss.
XX OS Hepatitis B virus.
XX FH Key Location/Qualifiers
XX FT CDS 10..669
XX FT /*tag= a
XX FT /product= "HBcAg"

```

```

FT FT /partial
XX FT /note= "No start codon"
PN WO2004035007-A2.
XX PD
XX PD 29-APR-2004.
XX PF 17-OCT-2003; 2003WO-US033178.
XX PR 17-OCT-2002; 2002US-0419279P.
XX PA (ORAG-) ORAGEN CORP.
XX PI Michaels F;
DR WPI; 2004-348329/32.
DR P-PSDB; ADO07221.
XX
XX Modulating a systemic immune response to a peptide in a mammal comprises
PT transmutosally administering a macromolecular aggregate of the peptide.
XX
XX Disclosure; SEQ ID NO 1; 81pp; English.
XX
CC The present sequence is the DNA sequence of the hepatitis B virus core
CC antigen (HBcAg) gene from HBV serotype ayw. A peptide comprising a HBV
CC protein can be used in claimed methods of the invention for modulating an
CC immune response in a mammal. A method of inducing a systemic immune
CC response to a peptide in a mammal comprises transmutosally administering
CC to the mammal a macromolecular aggregate of the peptide. The
CC macromolecular aggregate comprises at least 10 peptide subunits, may have
CC a molecular weight of over 1,000 kDa, and is preferably at least 5 nm in
CC diameter. It is resistant to digestive degradation, being stabilised in
CC aggregate form by chemical treatment and/or by recombinant protein
CC engineering of the peptide. The peptide preferably comprises a HBV
CC protein selected from HBV surface protein, nucleocapsid protein or
CC envelope protein. Transmucosal administration to a mammal of a
CC macromolecular aggregate of a HBV surface protein engenders a systemic
CC immune response in the mammal. A method of suppressing an immune response
CC in a mammal involves transmucosally administering a monomolecular peptide
CC that is resistant to digestive degradation and which may be stabilised by
CC chemical treatment or protein engineering, and which may be derived from
CC a HBV protein. A monomolecular peptide is useful for the induction of
CC oral tolerance when induction of systemic immunity is undesirable, e.g.
CC in cases of chronic infections.
XX
XX Sequence 669 BP; 155 A; 170 C; 148 G; 196 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 12; Length 669;
Best Local Similarity 85.0%; Pred. No. 3.3;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGTCGAUGUCCAUCC 20
DB 134 TAAGGGTCGATGCCATGCC 115

RESULT 12
AAV82691/C
ID AAV82691 standard; DNA; 1334 BP.
XX AC
XX AAV82691;
XX
XX 16-FEB-1999 (first entry)
XX
XX Fulminant hepatitis B virus genotype D variant FHBV12 sequence.
XX
XX Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
XX HBV-related disease; ss.
XX
XX Hepatitis B virus.
XX
XX WO9845421-A2.
XX

```

```

PD 15-OCT-1998.
XX
XX 08-APR-1998; 98WO-EF002048.
XX
XX 09-APR-1997; 97GB-00007221.
XX
XX (UNIU ) UNIV GLASGOW.
XX
XX Carman B;
XX
XX WPI; 1999-009329/01.
XX
XX New hepatitis B virus nucleic acid with combination of specific mutations
PT - useful for, e.g. detection of binding interactions between host or
PT viral proteins and HBV nucleic.
XX
XX Disclosure; Fig 5; 85pp; English.
XX
XX The present sequence represents part of the genome of a fulminant
CC Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC a mutation (i.e. alteration from the normal nucleotide in any of the
CC genotypes A to F) in at least two of the enhancer I region, the negative
CC regulatory element region, the enhancer II/ core upstream regulatory
CC sequence/ basal core promoter region, or a mutation which leads to an X-
CC peptide amino acid change to Cys or Met. The HBV variants of the
CC invention are used to detect binding interactions between host or viral
CC proteins and HBV nucleic acid. Probes that hybridise to any of the
CC specified mutated regions are used to detect HBV-related disease.
CC especially fulminant infection, but also severe chronic infection or
CC serologically unusual forms of disease. Combinations of the specified
CC mutations are associated with fulminant infections, probably because they
CC reduce the ability to bind inhibitory proteins in the host cell
XX
XX Sequence 1334 BP; 288 A; 363 C; 311 G; 372 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 20; DB 2; Length 1334;
Best Local Similarity 85.0%; Pred. No. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGTCGAUGUCCAUCC 20
DB 806 TAAGGGTCGATGCCATGCC 787

RESULT 13
AAV82688/C
ID AAV82688 standard; DNA; 1395 BP.
XX AC
XX AAV82688;
XX
XX 16-FEB-1999 (first entry)
XX
XX Fulminant hepatitis B virus genotype D variant FHBV5 sequence.
XX
XX Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
XX HBV-related disease; ss.
XX
XX Hepatitis B virus.
XX
XX WO9845421-A2.
XX
XX 15-OCT-1998.
XX
XX 08-APR-1998; 98WO-EF002048.
XX
XX 09-APR-1997; 97GB-00007221.
XX
XX (UNIU ) UNIV GLASGOW.
XX
XX Carman B;
XX
XX WPI; 1999-009329/01.
XX

```

XX New hepatitis B virus nucleic acid with combination of specific mutations
PT - useful for, e.g. detection of binding interactions between host or
PT viral proteins and HBV nucleic.
XX Disclosure; Fig 5; 85pp; English.
XX The present sequence represents part of the genome of a fulminant
CC Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC a mutation (i.e. alteration from the normal nucleotide in any of the
CC genotypes A to F) in at least two of the enhancer I region, the negative
CC regulatory element region, the enhancer II/ core upstream regulatory
CC sequence/ basal core promoter region, or a mutation which leads to an X-
CC peptide amino acid change to Cys or Met. The HBV variants of the
CC invention are used to detect binding interactions between host or viral
CC proteins and HBV nucleic acid. Probes that hybridise to any of the
CC specified mutated regions are used to detect HBV-related disease,
CC especially fulminant infection, but also severe chronic infection or
CC serologically unusual forms of disease. Combinations of the specified
CC mutations are associated with fulminant infections, probably because they
CC reduce the ability to bind inhibitory proteins in the host cell
XX Sequence 1395 BP; 277 A; 387 C; 331 G; 398 T; 0 U; 2 Other;
SQ Query Match 100.0%; Score 20; DB 2; Length 1395;
Best Local Similarity 85.0%; Pred. No. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAGGTCGAUGUCCAUGCC 20
Db |||||||:|:|:|
917 TAAGGTCGATGTCATGCC 898

RESULT 14
AAV82687/c
ID AAV82687 standard; DNA; 1400 BP.
XX AAV82687;
XX 16-FEB-1999 (first entry)
XX Fulminant hepatitis B virus genotype D variant FHBV4 sequence.
DE Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW HBV-related disease; ss.
XX Hepatitis B virus.
OS WO9845421-A2.
XX 15-OCT-1998.
XX 08-APR-1998; 98WO-EP002048.
XX 09-APR-1997; 97GB-00007221.
XX (UNIU) UNIV GLASGOW.
XX Carman B;
XX WPI; 1999-009329/01.
XX New hepatitis B virus nucleic acid with combination of specific mutations
PT - useful for, e.g. detection of binding interactions between host or
PT viral proteins and HBV nucleic.
XX Disclosure; Fig 5; 85pp; English.
XX The present sequence represents part of the genome of a fulminant
CC Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC a mutation (i.e. alteration from the normal nucleotide in any of the
CC genotypes A to F) in at least two of the enhancer I region, the negative
CC regulatory element region, the enhancer II/ core upstream regulatory
CC sequence/ basal core promoter region, or a mutation which leads to an X-
CC peptide amino acid change to Cys or Met. The HBV variants of the
CC invention are used to detect binding interactions between host or viral
CC proteins and HBV nucleic acid. Probes that hybridise to any of the
CC specified mutated regions are used to detect HBV-related disease,
CC especially fulminant infection, but also severe chronic infection or
CC serologically unusual forms of disease. Combinations of the specified
CC mutations are associated with fulminant infections, probably because they
CC reduce the ability to bind inhibitory proteins in the host cell
XX Sequence 1395 BP; 277 A; 387 C; 331 G; 398 T; 0 U; 2 Other;
SQ Query Match 100.0%; Score 20; DB 2; Length 1395;
Best Local Similarity 85.0%; Pred. No. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAGGTCGAUGUCCAUGCC 20
Db |||||||:|:|:|
917 TAAGGTCGATGTCATGCC 898

CC genotypes A to F) in at least two of the enhancer I region, the negative
CC regulatory element region, the enhancer II/ core upstream regulatory
CC sequence/ basal core promoter region, or a mutation which leads to an X-
CC peptide amino acid change to Cys or Met. The HBV variants of the
CC invention are used to detect binding interactions between host or viral
CC proteins and HBV nucleic acid. Probes that hybridise to any of the
CC specified mutated regions are used to detect HBV-related disease,
CC especially fulminant infection, but also severe chronic infection or
CC serologically unusual forms of disease. Combinations of the specified
CC mutations are associated with fulminant infections, probably because they
CC reduce the ability to bind inhibitory proteins in the host cell
XX Sequence 1400 BP; 287 A; 388 C; 332 G; 393 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 20; DB 2; Length 1400;
Best Local Similarity 85.0%; Pred. No. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAGGTCGAUGUCCAUGCC 20
Db |||||||:|:|:|
917 TAAGGTCGATGTCATGCC 898

RESULT 15
AAV82692/c
ID AAV82692 standard; DNA; 1445 BP.
XX AAV82692;
XX 16-FEB-1999 (first entry)
XX Fulminant hepatitis B virus genotype D variant FHBV13 sequence.
DE Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW HBV-related disease; ss.
XX Hepatitis B virus.
OS WO9845421-A2.
XX 15-OCT-1998.
XX 08-APR-1998; 98WO-EP002048.
XX 09-APR-1997; 97GB-00007221.
XX (UNIU) UNIV GLASGOW.
XX Carman B;
XX WPI; 1999-009329/01.
XX New hepatitis B virus nucleic acid with combination of specific mutations
PT - useful for, e.g. detection of binding interactions between host or
PT viral proteins and HBV nucleic.
XX Disclosure; Fig 5; 85pp; English.
XX The present sequence represents part of the genome of a fulminant
CC Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC a mutation (i.e. alteration from the normal nucleotide in any of the
CC genotypes A to F) in at least two of the enhancer I region, the negative
CC regulatory element region, the enhancer II/ core upstream regulatory
CC sequence/ basal core promoter region, or a mutation which leads to an X-
CC peptide amino acid change to Cys or Met. The HBV variants of the
CC invention are used to detect binding interactions between host or viral
CC proteins and HBV nucleic acid. Probes that hybridise to any of the
CC specified mutated regions are used to detect HBV-related disease,
CC especially fulminant infection, but also severe chronic infection or
CC serologically unusual forms of disease. Combinations of the specified
CC mutations are associated with fulminant infections, probably because they
CC reduce the ability to bind inhibitory proteins in the host cell

XX
SQ Sequence 1445 BP; 297 A; 406 C; 338 G; 404 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 1445;
Best Local Similarity 85.0%; Pred. No. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAAGGGTCGAUGUCCAUGCC 20
Db 917 TAAGGGTCGATGCCATGCC 898

Search completed: March 17, 2005, 06:48:43
Job time : 172.333 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-62

Perfect score: 20

Sequence: 1 taaggctgauguccaagcc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	95.0	686	4	BI954819
C 2	17.4	87.0	576	8	CC145068
C 3	17.4	87.0	763	9	CL248902
C 4	17	85.0	260	7	CN589115
C 5	17	85.0	451	5	BP643309
C 6	17	85.0	729	4	BJ705222
C 7	16.8	84.0	413	5	BP569691
C 8	16.8	84.0	431	9	CL260820
C 9	16.8	84.0	457	5	BP527907
C 10	16.8	84.0	509	9	TA230D01P
C 11	16.8	84.0	733	1	AI635930
C 12	16.8	84.0	773	8	BZ573222
C 13	16.8	84.0	1574	9	AG476230
C 14	16.4	82.0	272	2	BE922196
C 15	16.4	82.0	306	9	AG200800
C 16	16.4	82.0	356	9	CC836806
C 17	16.4	82.0	417	7	H70178
C 18	16.4	82.0	499	5	BQ116265
C 19	16.4	82.0	554	8	AQ433745
C 20	16.4	82.0	612	9	CL194160
C 21	16.4	82.0	622	5	BQ112668
C 22	16.4	82.0	633	9	CG461236
C 23	16.4	82.0	678	9	CG761335
C 24	16.4	82.0	684	9	CL369255

25	16.4	82.0	714	9	AG128210
26	16.4	82.0	772	8	BH257005
C 27	16.4	82.0	848	7	CO099982
28	16.4	82.0	851	4	BJ571152
29	16.4	82.0	924	9	CG289400
30	16.4	82.0	984	9	CG875596
31	16.4	82.0	1010	9	CC522611
C 32	16.4	82.0	1046	9	CL055525
33	16.4	82.0	1669	3	HSM800819
34	16.4	82.0	1669	3	HSM802167
35	16	80.0	168	4	BI052244
36	16	80.0	168	4	BI052309
37	16	80.0	532	5	EX433931
C 38	16	80.0	666	9	CG106546
C 39	16	80.0	698	8	BZ413874
C 40	16	80.0	740	9	CG235822
C 41	16	80.0	750	9	CG235810
C 42	16	80.0	810	9	CC716058
C 43	16	80.0	836	9	CG257272
44	16	80.0	963	9	CG302315
45	15.8	79.0	68	8	BH631190

ALIGNMENTS

RESULT 1
BI954819/c
LOCUS BI954819 686 bp mRNA linear EST 19-OCT-2001
DEFINITION HVSMEM0019017f Hordeum vulgare green seedling EST library
clone HVSMEM0019017f, mRNA sequence.
ACCESSION BI954819
VERSION BI954819.1 GI:16300646
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 686)
Wing, R., Close, T.J., Klein, A., Wise, R., Chin, A., Begum, D.,
Frisch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T.,
Simmons, J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Blumeria infected Morex (compatible) seedling
cDNA library
JOURNAL Unpublished (2001)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 417
Seq primer: AATTAACCTCACTAAGGG
High quality sequence start: 16
High quality sequence stop: 531.

FEATURES

source
1. .686
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMEM0019017f"
/cissue="type="green seedling leaf"
/lab_host="TJC121"
/clone_lib="Hordeum vulgare green seedling EST library
HVCDNA0014 (Blumeria infected)"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Morex (mla) plants were greenhouse grown in the R

Wise lab at Iowa State University, Ames, IA; 7 day old green seedlings were infected with isolate 5874 of *Blumeria graminis* f. sp. hordei, and leaves were harvested 24, 48 and 72 hr post-inoculation and snap frozen (Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Chin). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinbols A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. *Barley Genetics Newsletter* 31:29-30. (<http://wheat.ow.usda.gov/apacques/bam31/cover.html>)

```

CN589115.1  GI:47040917
EST.
KEYWORDS
SOURCE Tetrahymena thermophila
ORGANISM Tetrahymena thermophila
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Tetrahymenina; Tetrahymenidae; Tetrahymena.
1 (bases 1 to 260)
Garg, J., Pearlman, R.E. and Carlton, J.
PEPdbPub (http://ameobidia.bcm.umontreal.ca/public/pepdb/agrm.php)
Tetrahymena thermophila (TIGR)
Unpublished (2004)
JOURNAL
COMMENT Contact: PEPdb
Departement de Biochimie, Universite de Montreal
Email: pepdb-curator@bcm.umontreal.ca
Plate: 1398.

FEATURES
source
1..260
Location/Qualifiers
/organism="Tetrahymena thermophila"
/mol_type="mRNA"
/db_xref="taxon:5911"
/clone_lib="Normalized large"

ORIGIN
Query Match 85.0%; Score 17; DB 7; Length 260;
Best Local Similarity 82.4%; Pred. No. 5e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGTCGAUGUCCAUCC 20
|||||:|:|:|
DB 237 GGGTCGATGTCATGCC 221

RESULT 5
BP643309
LOCUS BP643309
DEFINITION BP643309.1 GI:49294779
ACCESSION BP643309
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 451)
Seki, M., Naruoka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,
Nakajima, M., Enju, A., Akiyama, K., Ono, Y., Muramatsu, M.,
Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)
21932900
JOURNAL
MEDLINE
PUBMED
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
reversed clone; Please visit our web site
(http://pfweb.gsc.riken.go.jp/) for further details.

FEATURES
source
1..451
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL19-61-I08"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAFL19"
/note="Site 1: BanHI; Site 2: SalI; Subtraction Library"

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

ORIGIN
Query Match 85.0%; Score 17; DB 5; Length 451;
Best Local Similarity 82.4%; Pred. No. 6.5e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 AGGTCGAUGUCCAUCC 19
|||||:|:|:|
DB 431 AGGTCGATGTCATGC 447

RESULT 6
BJ705222/c
LOCUS BJ705222
DEFINITION BJ705222 MF01FFA cDNA Oryzias latipes cdna clone MF01FFA013all 5',
mRNA sequence.
ACCESSION BJ705222
VERSION BJ705222.1 GI:45246102
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 729)
Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
Medaka EST Project in Takeda's lab
Unpublished (2001)
JOURNAL
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tschini@genes.nig.ac.jp.

FEATURES
source
1..729
Location/Qualifiers
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="hd-r8"
/db_xref="taxon:8090"
/clone="MF01FFA013all"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_lib="MF01FFA CDNA"

ORIGIN
Query Match 85.0%; Score 17; DB 4; Length 729;
Best Local Similarity 82.4%; Pred. No. 7e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAGGTCGAUGUCCAUCC 18
|||||:|:|:|
DB 182 AAGGTCGATGTCATGC 166

RESULT 7
BP569691
LOCUS BP569691
DEFINITION BP569691 RAFL14 Arabidopsis thaliana cdna clone RAFL14-68-M24 3',
mRNA sequence.
ACCESSION BP569691
VERSION BP569691.1 GI:48985457
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 413)
Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,

```

```

CN589115.1  GI:47040917
EST.
KEYWORDS
SOURCE Tetrahymena thermophila
ORGANISM Tetrahymena thermophila
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Tetrahymenina; Tetrahymenidae; Tetrahymena.
1 (bases 1 to 260)
Garg, J., Pearlman, R.E. and Carlton, J.
PEPdbPub (http://ameobidia.bcm.umontreal.ca/public/pepdb/agrm.php)
Tetrahymena thermophila (TIGR)
Unpublished (2004)
JOURNAL
COMMENT Contact: PEPdb
Departement de Biochimie, Universite de Montreal
Email: pepdb-curator@bcm.umontreal.ca
Plate: 1398.

FEATURES
source
1..260
Location/Qualifiers
/organism="Tetrahymena thermophila"
/mol_type="mRNA"
/db_xref="taxon:5911"
/clone_lib="Normalized large"

ORIGIN
Query Match 85.0%; Score 17; DB 7; Length 260;
Best Local Similarity 82.4%; Pred. No. 5e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGTCGAUGUCCAUCC 20
|||||:|:|:|
DB 237 GGGTCGATGTCATGCC 221

RESULT 5
BP643309
LOCUS BP643309
DEFINITION BP643309.1 GI:49294779
ACCESSION BP643309
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 451)
Seki, M., Naruoka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,
Nakajima, M., Enju, A., Akiyama, K., Ono, Y., Muramatsu, M.,
Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)
21932900
JOURNAL
MEDLINE
PUBMED
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
reversed clone; Please visit our web site
(http://pfweb.gsc.riken.go.jp/) for further details.

FEATURES
source
1..451
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL19-61-I08"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAFL19"
/note="Site 1: BanHI; Site 2: SalI; Subtraction Library"

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

ORIGIN
Query Match 85.0%; Score 17; DB 5; Length 451;
Best Local Similarity 82.4%; Pred. No. 6.5e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 AGGTCGAUGUCCAUCC 19
|||||:|:|:|
DB 431 AGGTCGATGTCATGC 447

RESULT 6
BJ705222/c
LOCUS BJ705222
DEFINITION BJ705222 MF01FFA cDNA Oryzias latipes cdna clone MF01FFA013all 5',
mRNA sequence.
ACCESSION BJ705222
VERSION BJ705222.1 GI:45246102
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 729)
Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
Medaka EST Project in Takeda's lab
Unpublished (2001)
JOURNAL
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tschini@genes.nig.ac.jp.

FEATURES
source
1..729
Location/Qualifiers
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="hd-r8"
/db_xref="taxon:8090"
/clone="MF01FFA013all"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_lib="MF01FFA CDNA"

ORIGIN
Query Match 85.0%; Score 17; DB 4; Length 729;
Best Local Similarity 82.4%; Pred. No. 7e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAGGTCGAUGUCCAUCC 18
|||||:|:|:|
DB 182 AAGGTCGATGTCATGC 166

RESULT 7
BP569691
LOCUS BP569691
DEFINITION BP569691 RAFL14 Arabidopsis thaliana cdna clone RAFL14-68-M24 3',
mRNA sequence.
ACCESSION BP569691
VERSION BP569691.1 GI:48985457
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 413)
Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,

```

Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M., Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)

TITLE
MEDLINE
JOURNAL
PUBMED
COMMENT

Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060

Email: msekic@rken.go.jp
reversed clone; Please visit our web site
(http://pfweb.gsc.riken.go.jp/) for further details.

FEATURES
source

1. 413
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL14-68-M24"
/tissue_type="root"
/lab_host="DH10B"
/clone_lib="RAFL14"
/notes="Site_1: BamHI; Site_2: SalI"

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 413;
Best Local Similarity 75.0%; Pred. No. 8.2e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCCAUCC 20
|||||:|:|:|:|:
Db 383 TAAGGTCGATGTCATGAC 402

RESULT 8

CL260820 431 bp DNA linear GSS 02-FEB-2004
ZMMBB0619B24r ZMMBBB (HindIII) Zea mays genomic clone
ZMMBB0619B24 3', genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS

Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wang,R. and Messing,J.
Sequencing of the maize genome at PGIR (2003c)

TITLE
JOURNAL
COMMENT

Unpublished (2003)
Contact: Bharti,A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801

Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends

High quality sequence start: 116.
Location/Qualifiers

FEATURES
source

1. 431
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBB0619B24"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 431;
Best Local Similarity 85.0%; Pred. No. 8.3e+02;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCCAUCC 20
|||||:|:|:|:|:
Db 328 TAAGGTCGATGCCAGGCC 347

RESULT 9

BP527907 457 bp mRNA linear EST 28-SBP-2004
BP527907 MAT001 Nicotiana tabacum cDNA clone BV12728, mRNA
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BP527907.1 GI:52831634
Nicotiana tabacum (common tobacco)
Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE

1 (bases 1 to 457)
Matsuoka,K., Tashiro,G., Horiguchi,T., Demura,T. and Fukuda,H.
Profiling growth-phase dependent gene expression of tobacco BY-2
cells by comprehensive microarray analysis
Unpublished (2003)

JOURNAL
COMMENT

Contact: Ken Matsuoka
Morphogenesis Research Group
RIKEN Plant Science Center
1-7-2 Suehirocho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9575
Fax: 81-45-503-9573

Email: by2@psc.riken.go.jp, URL: http://mrq.psc.riken.go.jp/strc/
The cDNA library was constructed from mRNA isolated from lag (9 h),
log (72 h) and stationary (7 days) old BY-2 cells.
Seq primer: M13 forward.

FEATURES
source

1. 457
Location/Qualifiers
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/cultivar="Bright Yellow No.2"
/db_xref="taxon:4097"
/clone="BY12728"
/cell_line="BY-2"
/clone_lib="MAT001"
/note="Vector: pGEM-T easy; primer: M13 forward; mRNA
obtained from lag, log and stationary phase cells"

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 457;
Best Local Similarity 75.0%; Pred. No. 8.3e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGTCGAUGUCCAUCC 20
|||||:|:|:|:|:
Db 297 TAAGGTCGATGCCATGCC 316

RESULT 10

TA230D01P 509 bp DNA linear GSS 13-DEC-2000
T. brucei sheared genomic DNA clone 230d01, forward sequence,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

AL481016
GI:11846785
GSS.
Trypanosoma brucei

```

ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 509)
AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
Melville, S.B., Rajandream, M.A. and Barrell, B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
rhl@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 TUTat 10.1) was mechanically sheared
to give a tight size distribution (
4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES
source
1..509
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="230d01"
ORIGIN
Query Match 84.0%; Score 16.8; DB 9; Length 509;
Best Local Similarity 80.0%; Pred. No. 8.5e+02;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGTCGAUGUCCAUGCC 20
|||||:|:|:|:|:|:|
Db 355 TAAGGGTTGATGCCAGGCC 374

RESULT 11
AI635930
LOCUS 733 bp mRNA linear EST 16-DEC-1999
DEFINITION tz82c11.x1 NCI-CGAP Panel Homo sapiens cDNA clone IMAGE:2295092 3'
similar to gb:J03490 DIHYDROLIPOAMIDE DEHYDROGENASE PRECURSOR
(HUMAN);, mRNA sequence.
ACCESSION AI635930
VERSION AI635930.1 GI:4687260
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 733)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-rc@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1107 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 401.
FEATURES
source
1..733
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2295092"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Panel"
/notes="Organ: Pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"
ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 733;
Best Local Similarity 75.0%; Pred. No. 8.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGTCGAUGUCCAUGCC 20
|||||:|:|:|:|:|:|
Db 557 TAAGGGTCGATGTCATGAC 576

RESULT 12
BZ573222/c
LOCUS 773 bp DNA linear GSS 17-DEC-2002
DEFINITION msh2_3006.y2 msh Pseudomonas aeruginosa genomic clone msh2_3006,
genomic survey sequence.
ACCESSION BZ573222
VERSION BZ573222.1 GI:27208283
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 773)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
1..773
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msh2_3006"
/clone_lib="msh"
/notes="Environmental isolate. Whole genomic shotgun
library."
ORIGIN
Query Match 84.0%; Score 16.8; DB 8; Length 773;
Best Local Similarity 80.0%; Pred. No. 9e+02;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGTCGAUGUCCAUGCC 20
|||||:|:|:|:|:|:|
Db 294 TAATGGTCGATGCCAGGCC 275

RESULT 13
AG476230/c
LOCUS 1574 bp DNA linear GSS 04-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-369C21.TJ, genomic survey
sequence.

```

```

ACCESSION AG476230
VERSION AG476230.1 GI:48183460
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS BAC end Sequences of Library MSMg01
TITLE Unpublished
JOURNAL
REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS BAC end Sequences of Library MSMg01
TITLE Unpublished
JOURNAL
COMMENT 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@isc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Teukuba Institute, Bio Resource Center.
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : ECORI.
R.Site 2 : ECORI.
FEATURES
source
1. .1574
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-369C21.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match 84.0%; Score 16.8; DB 9; Length 1574;
Best Local Similarity 75.0%; Pred. No. 1e+03;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 TAAGGTCGAUGUCCAUCC 20
Db 300 TATGGTCGATGTCATCCC 281
RESULT 14
BE922196 272 bp mRNA linear EST 07-MAR-2003
LOCUS EST425953 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION cSTB18G4 5' sequence, mRNA sequence.
ACCESSION BE922196
VERSION BE922196.1 GI:10448260
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
REFERENCE van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
AUTHORS Rutterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R.,
Kroning,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/.
FEATURES
source
1. .272
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB18G4"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/clone_lib="potato leaves and petioles"
/notes="Vector: pBlueScript SK(-); Site 1: EcorI; Site 2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
ORIGIN
Query Match 82.0%; Score 16.4; DB 2; Length 272;
Best Local Similarity 77.8%; Pred. No. 1.2e+03;
Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 AGGTCGAUGUCCAUCC 20
Db 29 AGGTCGATGTCATGCC 46
RESULT 15
AG200800/c 306 bp DNA linear GSS 06-MAR-2004
LOCUS Pan troglodytes DNA, clone: RP43-082P16.T7, genomic survey
DEFINITION sequence.
ACCESSION AG200800
VERSION AG200800.1 GI:45232975
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
AUTHORS Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
TITLE BAC end sequences of Library RP-43
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 306)
AUTHORS Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
TITLE Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
JOURNAL Direct Submission
COMMENT Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
52, Oun-dong, Yuseong-gu, Daejeon 305-333, Korea
(E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
Tel:82-42-866-7181, Fax:82-42-860-4409)
Clones are derived from the chimpanzee BAC library RP-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: T7
LIBRARY
Vector : pBACe3.6
R.Site 1 : ECORI.
R.Site 2 : ECORI.
FEATURES
source
1. .306
/organism="Pan troglodytes"
/mol_type="genomic DNA"

```

```

/db_xref="taxon:9598"
/clone="RP43-082P16.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RP-43 Chimpanzee Male BAC Library"

```

ORIGIN

```

Query Match      82.0%; Score 16.4; DB 9; Length 306;
Best Local Similarity 77.8%; Pred. NO. 1.3e+03;
Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 TAAGGTCGAUGUCCAUG 18
         |||||  ||:|:|:|
Db      35 TAAGGTAGATGCCAIG 18

```

```

Search completed: March 17, 2005, 11:07:49
Job time : 1389.27 secs

```

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-63

Perfect score: 20

Sequence: 1 taaggguccgauguccatgcc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6 AR027817	AR027817 Sequence
2	20	100.0	30	6 AR027842	AR027842 Sequence
3	20	100.0	30	6 AR027843	AR027843 Sequence
C 4	20	100.0	81	6 I92348	I92348 Sequence 9
C 5	20	100.0	174	14 S77749	S77749 preC {X/pre
C 6	20	100.0	253	14 AY329529	AY329529 Hepatitis
C 7	20	100.0	253	14 AY329562	AY329562 Hepatitis
C 8	20	100.0	253	14 AY329568	AY329568 Hepatitis
C 9	20	100.0	253	14 AY329573	AY329573 Hepatitis
C 10	20	100.0	253	14 AY329581	AY329581 Hepatitis
C 11	20	100.0	294	14 AF390000	AF390000 Hepatitis
C 12	20	100.0	333	14 HPHBED	L12359 Hepatitis B
C 13	20	100.0	398	14 AB167603	AB167603 Hepatitis
C 14	20	100.0	398	14 AB167637	AB167637 Hepatitis
C 15	20	100.0	406	14 AB163815	AB163815 Hepatitis
C 16	20	100.0	406	14 AB163817	AB163817 Hepatitis
C 17	20	100.0	439	14 AY254503	AY254503 Hepatitis
C 18	20	100.0	488	14 AY274419	AY274419 Hepatitis
C 19	20	100.0	488	14 AY274420	AY274420 Hepatitis

C 20	20	100.0	488	14	AY274422	AY274422 Hepatitis
C 21	20	100.0	488	14	AY274427	AY274427 Hepatitis
C 22	20	100.0	488	14	AY274428	AY274428 Hepatitis
C 23	20	100.0	488	14	AY274429	AY274429 Hepatitis
C 24	20	100.0	488	14	AY274430	AY274430 Hepatitis
C 25	20	100.0	488	14	AY274431	AY274431 Hepatitis
C 26	20	100.0	488	14	AY274432	AY274432 Hepatitis
C 27	20	100.0	488	14	AY274433	AY274433 Hepatitis
C 28	20	100.0	488	14	AY274434	AY274434 Hepatitis
C 29	20	100.0	488	14	AY274436	AY274436 Hepatitis
C 30	20	100.0	493	14	S79556	S79556 X, preC [he
C 31	20	100.0	548	14	AY382500	AY382500 Hepatitis
C 32	20	100.0	548	14	AY382501	AY382501 Hepatitis
C 33	20	100.0	548	14	AY382502	AY382502 Hepatitis
C 34	20	100.0	548	14	AY382521	AY382521 Hepatitis
C 35	20	100.0	548	14	AY382522	AY382522 Hepatitis
C 36	20	100.0	548	14	AY382523	AY382523 Hepatitis
C 37	20	100.0	548	14	AY382524	AY382524 Hepatitis
C 38	20	100.0	548	14	AY382525	AY382525 Hepatitis
C 39	20	100.0	548	14	AY382526	AY382526 Hepatitis
C 40	20	100.0	548	14	AY382527	AY382527 Hepatitis
C 41	20	100.0	609	14	AF289954	AF289954 Hepatitis
C 42	20	100.0	609	14	AF289965	AF289965 Hepatitis
C 43	20	100.0	626	14	AY254500	AY254500 Hepatitis
C 44	20	100.0	639	6	AX278066	AX278066 Sequence
C 45	20	100.0	639	6	AX342485	AX342485 Sequence

ALIGNMENTS

RESULT 1
LOCUS AR027817 20 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 15 from patent US 5856459.
ACCESSION AR027817
VERSION AR027817.1 GI:5938637
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 20)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and
Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 15 05-JAN-1999;
FEATURES Location/Qualifiers
source 1..20
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 85.0%; Pred. No. 14;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAGGGUCCGAUGUCCATGCC 20
|||||:||||:|||||
Db 1 TAAGGGTCGATGTCATGCC 20

RESULT 2
LOCUS AR027842 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 40 from patent US 5856459.
ACCESSION AR027842
VERSION AR027842.1 GI:5938662
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 30)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

```

Mills, J.S.
TITLE      Oligonucleotides specific for hepatitis B virus
JOURNAL    Patent: US 5856459-A 40 05-JAN-1999;
FEATURES   Location/Qualifiers
            source
            1..30
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 85.0%; Pred. No. 14;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20
Db 11 TAAGGGTCGATGTCATGCC 30

RESULT 3
AR027843
LOCUS      AR027843          30 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION Sequence 41 from patent US 5856459.
ACCESSION  AR027843
VERSION     AR027843.1  GI:5938663
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 30)
AUTHORS    Frank, B.L., Roberts, P.C., Goodchild, J., Craig, J., Charles. and
            Mills, J.S.
TITLE      Oligonucleotides specific for hepatitis B virus
JOURNAL    Patent: US 5856459-A 41 05-JAN-1999;
FEATURES   Location/Qualifiers
            source
            1..30
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 85.0%; Pred. No. 14;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20
Db 11 TAAGGGTCGATGTCATGCC 30

RESULT 4
I92348/c
LOCUS      I92348          81 bp      DNA      linear      PAT 01-DEC-1998
DEFINITION Sequence 9 from patent US 5728518.
ACCESSION  I92348
VERSION     I92348.1  GI:3936818
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 81)
AUTHORS    Carmichael, E.
TITLE      Antiviral poly- and oligonucleotides
JOURNAL    Patent: US 5728518-A 9 17-MAR-1998;
FEATURES   Location/Qualifiers
            source
            1..81
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 81;
Best Local Similarity 85.0%; Pred. No. 15;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20

```

```

|||||:||||:|||||
71 TAAGGGTCGATGTCATGCC 52

RESULT 5
S77749/c
LOCUS      S77749          174 bp      DNA      linear      VRL 06-MAY-2003
DEFINITION preC (X/preC region, deletion mutant) [hepatitis B virus HBV,
            host-human, serum, patient 5 isolate, Genomic DNA Mutant, 174 nt].
ACCESSION  S77749
VERSION     S77749.1  GI:999129
KEYWORDS
SOURCE      Hepatitis B virus
ORGANISM    Hepatitis B virus
REFERENCE   1 (bases 1 to 174)
AUTHORS    Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
TITLE      1 (bases 1 to 174)
JOURNAL    Feitelson, M.A., Duan, L.X., Guo, J., and Blumberg, B.S.
MEDLINE    X region deletion mutants associated with surface antigen-positive
PUBMED     hepatitis B virus infections
REMARK     Gastroenterology 108 (6), 1810-1819 (1995)
           95285997
           7768387
Gensbank staff at the National Library of Medicine created this
entry [NCBI gibbsq 165980] from the original journal article.
FEATURES   Location/Qualifiers
            source
            1..174
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10407"
            69..>174
            /gene="preC"
            /note="no start codon found"
gene
ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 174;
Best Local Similarity 85.0%; Pred. No. 16;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20
Db 172 TAAGGGTCGATGTCATGCC 153

RESULT 6
AY329529/c
LOCUS      AY329529          253 bp      DNA      linear      VRL 08-JUN-2004
DEFINITION Hepatitis B virus isolate A611252E X protein gene, partial cds; and
            preC/C protein gene, complete cds.
ACCESSION  AY329529
VERSION     AY329529.1  GI:37625315
KEYWORDS
SOURCE      Hepatitis B virus
ORGANISM    Hepatitis B virus
REFERENCE   1 (bases 1 to 253)
AUTHORS    Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
TITLE      1 (bases 1 to 253)
JOURNAL    Sitnik, R., Rebello Pinho, J.R., Bertolini, D.A., Bernardini, A.P., Da
PUBMED     Silva, L.C. and Carrilho, F.J.
REMARK     Hepatitis B Virus Genotypes and Precore and Core Mutants in
           Brazilian Patients
           J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
           15184419
           2 (bases 1 to 253)
           Rebell Pinho, J.R., Sitnik, R., Carrilho, F.J., Da Silva, L.C. and
           Bernardini, A.P.
           Direct Submission
           Submitted (23-JUN-2003) Research & Development, Laboratorio
           Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
           01402-001, Brazil
           Location/Qualifiers
           source
           1..253
           /organism="Hepatitis B virus"
           /mol_type="genomic DNA"
           /isolate="A611252E"

```

```

/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95961.1"
/db_xref="GI:37625316"
/translation="STTDLEAYFKDCLFKDWEELBEIRLMIFVLGGCRHKLVCPAP
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95962.1"
/db_xref="GI:37625317"
/translation="MQLFHLCLVISCTPFOASKLCLGLW"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20
|||||:||||:|||||
Db 237 TAAGGTCGATGCCATGCC 218

RESULT 7
AV329562/c
LOCUS
DEFINITION
Hepatitis B virus isolate D273984E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329562
VERSION
AY329562.1 GI:37625413
KEYWORDS
.
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
15184419
PUBMED
2 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratorio
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
10402-001, Brazil

FEATURES
Location/Qualifiers
source
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D273984E"
/db_xref="taxon:10407"
<1..158
/product="X protein"
/protein_id="AAQ95961.1"
/db_xref="GI:37625432"
/translation="STTDLEAYFKDCLFKDWEELBEIRLMIFVLGGCRHKLVCPAP
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95962.1"
/db_xref="GI:37625414"
/translation="STTDLEAYFKDCLFKDWEELBEIRLMIFVLGGCRHKLVCPAP
CNFF TSA"
134..217
/codon_start=3
/product="X protein"
/protein_id="AAQ95961.1"
/db_xref="GI:37625415"
/translation="MQLFHLCLVISCTPFOASKLCLGLW"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20
|||||:||||:|||||
Db 237 TAAGGTCGATGCCATGCC 218

RESULT 8
AV329568/c
LOCUS
DEFINITION
Hepatitis B virus isolate D296668E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329568
VERSION
AY329568.1 GI:37625431
KEYWORDS
.
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
15184419
PUBMED
2 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratorio
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
10402-001, Brazil

FEATURES
Location/Qualifiers
source
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D296668E"
/db_xref="taxon:10407"
<1..158
/product="X protein"
/protein_id="AAQ95938.1"
/db_xref="GI:37625432"
/translation="STTDLEAYFKDCLFKDWEELBEIRLMIFVLGGCRHKLVCPAP
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95939.1"
/db_xref="GI:37625433"
/translation="MQLFHLCLVISCTPFOASKLCLGLW"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20
|||||:||||:|||||
Db 237 TAAGGTCGATGCCATGCC 218

RESULT 9
AV329573/c
LOCUS
DEFINITION
Hepatitis B virus isolate D804917E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329573
VERSION
AY329573.1 GI:37625446
KEYWORDS
.
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

```

```

/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95961.1"
/db_xref="GI:37625316"
/translation="STTDLEAYFKDCLFKDWEELBEIRLMIFVLGGCRHKLVCPAP
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95962.1"
/db_xref="GI:37625317"
/translation="MQLFHLCLVISCTPFOASKLCLGLW"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20
|||||:||||:|||||
Db 237 TAAGGTCGATGCCATGCC 218

RESULT 8
AV329568/c
LOCUS
DEFINITION
Hepatitis B virus isolate D296668E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329568
VERSION
AY329568.1 GI:37625431
KEYWORDS
.
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
15184419
PUBMED
2 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratorio
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
10402-001, Brazil

FEATURES
Location/Qualifiers
source
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D296668E"
/db_xref="taxon:10407"
<1..158
/product="X protein"
/protein_id="AAQ95938.1"
/db_xref="GI:37625432"
/translation="STTDLEAYFKDCLFKDWEELBEIRLMIFVLGGCRHKLVCPAP
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95939.1"
/db_xref="GI:37625433"
/translation="MQLFHLCLVISCTPFOASKLCLGLW"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 17;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20
|||||:||||:|||||
Db 237 TAAGGTCGATGCCATGCC 218

RESULT 9
AV329573/c
LOCUS
DEFINITION
Hepatitis B virus isolate D804917E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329573
VERSION
AY329573.1 GI:37625446
KEYWORDS
.
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

```

CDS		/isolate="D639472E" /db_xref="taxon:10407" <1..158 /codon_start=3 /product="X protein" /protein_id="AAQ95964.1" /db_xref="GI:37625471" /translation="STTDLEAYFKDCLFKDWBELGEBELLLIFVLGGCRHKLVCAPAF CNFFTSA" 134..217 /codon_start=1 /product="preC/C protein" /protein_id="AAQ95965.1" /db_xref="GI:37625472" /translation="MQLFHLCLIIISCSCTVQASKLCLGWL"	
CDS		100.0%; Score 20; DB 14; Length 253; Best Local Similarity 85.0%; Pred. No. 17; Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;	
ORIGIN			
QY	1	TAAGGUGCAUGUCCATGCC 20	
DB	237	: : TAAGGTCGATGTCATGCC 218	
RESULT 11			
AF390000/c		294 bp DNA linear VRL 06-MAR-2002	
LOCUS		Hepatitis B virus isolate D3 X protein gene, partial cds; and	
DEFINITION		nonfunctional precore/core protein gene, partial sequence.	
ACCESSION		AF390000	
VERSION		AF390000.1 GI:16266099	
KEYWORDS			
SOURCE		Hepatitis B virus	
ORGANISM		Hepatitis B virus	
REFERENCE		Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.	
AUTHORS		Castro,L.D., Niel,C. and Gomes,S.A.	
TITLE		Low frequency of mutations in the core promoter and precore regions of hepatitis B virus in anti-HBe positive Brazilian carriers	
JOURNAL		BMC Microbiol. 1 (1), 10 (2001)	
PUBMED		11472634	
REFERENCE		2 (bases 1 to 294)	
AUTHORS		De Castro,L., Niel,C. and Gomes,S.A.	
TITLE		Direct Submission	
JOURNAL		Submitted (11-JUN-2001) Virology, FIOCRUZ, Av. Brasil 4365, Rio de Janeiro, RJ 21045-900, Brazil	
FEATURES		Location/Qualifiers	
source		1..294 /organism="Hepatitis B virus" /mol_type="genomic DNA" /isolate="D3" /db_xref="taxon:10407" <1..119 /codon_start=3 /product="X protein" /protein_id="AAL16752.1" /db_xref="GI:16266100" /translation="FKDWELGDDSLMIYVLGGCRHKLVCAPAFCNFTSA" 95..>294 /note="nonfunctional precore/core protein due to mutation"	
CDS		100.0%; Score 20; DB 14; Length 294; Best Local Similarity 85.0%; Pred. No. 17; Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;	
ORIGIN			
QY	1	TAAGGUGCAUGUCCATGCC 20	
DB	198	: : TAAGGTCGATGTCATGCC 179	

[illegible]

Job time : 683.733 secs

Best Local Similarity 85.0%; Pred. No. 18;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGUGCAUCCATGCC 20
|||||:|||||
Db 287 TAAGGTCGATGTCATGCC 268

RESULT 15
AB163815/c
LOCUS AB163815 406 bp DNA linear VRL 29-JUN-2004
DEFINITION Hepatitis B virus X, prec/C genes for polyproteins, isolate:
SAF662.
ACCESSION AB163815
VERSION AB163815.1 GI:49387444
KEYWORDS
SOURCE Hepatitis B virus
ORGANISM Hepatitis B virus
REFERENCE 1
Hasegawa, I., Tanaka, Y., Kramvis, A., Kato, T., Suganuchi, F.,
Acharya, S. K., Orito, E., Ueda, R., Kew, M. C. and Mizokami, M.
TITLE Novel hepatitis B virus genotype a subtyping assay that
distinguishes subtype aa from ae and its application in
epidemiological studies
J. Virol. 78 (14), 7575-7581 (2004)
JOURNAL
PUBMED 15220432
REFERENCE 2 (bases 1 to 406)
Hasegawa, I., Tanaka, Y. and Mizokami, M.
AUTHORS Direct Submission
TITLE Submitted (24-FEB-2004) Izumi Hasegawa, Nagoya City University
JOURNAL Graduate School, Department of Internal Medicine and Molecular
Science; 1 Kawasaki, Mizuho-cho, Mizuho-ku, Nagoya, Aichi 467-8601,
Japan (E-mail: izu-hase@med.nagoya-cu.ac.jp, Tel: 81-52-853-8216,
Fax: 81-52-852-0849)

FEATURES
source
1..406
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="SAF662"
/db_xref="taxon:10407"
1..208
/gene="X"
<1..208
/gene="X"
/codon_start=2
/product="polyprotein"
/protein_id="BAD24924.1"
/db_xref="GI:49387445"
/translation="QILPKVLHKRTLGLPAMSTTDLEAYFKDCVFKDWBELGETRLM
IFVLGGRHKLVCAPSSCNFTSA"
184..406
/gene="prec/C"
184..>406
/gene="prec/C"
/codon_start=1
/product="polyprotein"
/protein_id="BAD24925.1"
/db_xref="GI:49387446"
/translation="MQLFHLCLLIISCTPTQASKICLGWLNGMDIDPYKEGATVEL
LSFLPQDFPFSVRDLDTASALYREALSP"

gene
CDS
184..>406
/gene="prec/C"
/codon_start=1
/product="polyprotein"
/protein_id="BAD24925.1"
/db_xref="GI:49387446"
/translation="MQLFHLCLLIISCTPTQASKICLGWLNGMDIDPYKEGATVEL
LSFLPQDFPFSVRDLDTASALYREALSP"

gene
CDS
184..>406
/gene="prec/C"
/codon_start=1
/product="polyprotein"
/protein_id="BAD24925.1"
/db_xref="GI:49387446"
/translation="MQLFHLCLLIISCTPTQASKICLGWLNGMDIDPYKEGATVEL
LSFLPQDFPFSVRDLDTASALYREALSP"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 406;
Best Local Similarity 85.0%; Pred. No. 18;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGUGCAUCCATGCC 20
|||||:|||||
Db 287 TAAGGTCGATGTCATGCC 268

Search completed: March 17, 2005, 08:14:17

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612a-63
Perfect score: 20
Sequence: 1 taagggucauguccatgcc 20

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	2	AAT72572 Hepatitis
2	20	100.0	20	2	AAT72573 Hepatitis
3	20	100.0	30	2	AAT72619 Hepatitis
4	20	100.0	30	2	AAT72620 Hepatitis
5	20	100.0	30	2	AAT72618 Hepatitis
6	20	100.0	30	2	AAT72621 Hepatitis
C 7	20	100.0	31	10	Adc64743
C 8	20	100.0	639	6	AAD27422
C 9	20	100.0	639	6	AAD31509
C 10	20	100.0	663	3	AAA71734
C 11	20	100.0	669	12	ADO07220
C 12	20	100.0	1334	2	AAV82691
C 13	20	100.0	1395	2	AAV82688
C 14	20	100.0	1400	2	AAV82687
C 15	20	100.0	1445	2	AAV82692
C 16	20	100.0	1445	2	AAV82685
C 17	20	100.0	1445	2	AAV82690
C 18	20	100.0	1445	2	AAV82684
C 19	20	100.0	1500	2	AAV82686
C 20	20	100.0	1500	2	AAV82689

C 21	20	100.0	2342	1	AA93072
C 22	20	100.0	3182	6	AA93175
C 23	20	100.0	3182	9	ACA6242
C 24	20	100.0	3182	10	AA60866
C 25	20	100.0	5618	2	AAQ88310
C 26	20	100.0	7991	6	AA516094
C 27	20	100.0	8007	6	AA516092
C 28	20	100.0	8717	6	AA516093
C 29	19	95.0	34	10	ADJ94544
C 30	19	95.0	34	10	ADJ94545
C 31	19	95.0	39	13	ADR89273
C 32	19	95.0	39	13	ADR89266
C 33	18.8	94.0	30	2	AAQ45813
C 34	18.8	94.0	30	2	AAV07810
C 35	18.8	94.0	30	2	AAV08309
C 36	18.8	94.0	50	2	AAQ06723
C 37	18.4	92.0	22	10	ADG46961
C 38	18.4	92.0	22	11	ADM83206
C 39	18.4	92.0	24	6	ABK44212
C 40	18.4	92.0	24	6	ABK67439
C 41	18.4	92.0	24	6	ABK67506
C 42	18.4	92.0	24	10	ADE10976
C 43	18.4	92.0	24	10	ADE10976
C 44	18.4	92.0	24	10	ADG46965
C 45	18.4	92.0	24	11	ADM83210

ALIGNMENTS

RESULT 1
AAT72572
ID AAT72572 standard; DNA; 20 BP.
XX
AC AAT72572;
XX
DT 03-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV92b.
XX
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..20
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"

WO9639502-A1.

12-DEC-1996.

04-JUN-1996; 96WO-BP002432.

06-JUN-1995; 95US-00467397.

(HOFF) HOFFMANN LA ROCHE & CO AG F.
(HYBR-) HYBRIDON INC.

Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
Roberts NA, Roberts PC, Slade A;
WPI; 1997-043124/04.

Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
used in the detection and treatment of HBV infection.

Claim 1; Page 12; 81pp; English.

The present sequence represents a synthetic oligonucleotide HBV92b which
is complementary to a portion of the hepatitis B virus (HBV) RNA. The
antisense oligonucleotide may be used to detect the presence of HBV in a

CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non- contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection

XX SQ Sequence 20 BP; 4 A; 5 C; 6 G; 5 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 85.0%; Pred. No. 2.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGCAUGCCATGCC 20
|||||:|||||
Db 1 TAAGGGTCGATGCCATGCC 20

RESULT 2
AAT72573
ID AAT72573 standard; DNA; 20 BP.
XX AC AAT72573;
AC AAT72573;
DT 03-SEP-1997 (first entry)
XX DE Hepatitis B virus RNA antisense oligonucleotide HBV92Mb.
XX HBV; HBV infection; inhibition; replication; ss.
XX Synthetic.

Key	Location/Qualifiers
FT misc_feature	1..20
FT misc_RNA	11..20
FT modified_base	11
FT modified_base	12
FT modified_base	13
FT modified_base	14
FT modified_base	15
FT modified_base	16
FT modified_base	17
FT modified_base	18
FT modified_base	19
FT modified_base	20

/tag= a
 /note= "Internucleotide linkages are phosphorothioate"
 /tag= b
 /note= "2'-O-Me RNA"
 /tag= c
 /mod_base= um
 /tag= d
 /mod_base= gm
 /tag= e
 /mod_base= um
 /tag= f
 /mod_base= cm
 /tag= g
 /mod_base= cm
 /tag= h
 /mod_base= OTHER
 /note= "2'-O-methyladenosine"
 /tag= i
 /mod_base= um
 /tag= j
 /mod_base= gm
 /tag= k
 /mod_base= cm
 /tag= l
 /mod_base= cm

XX W09639502-A1.
XX 12-DEC-1996.

PF 04-JUN-1996; 96WO-EP002432.
XX 06-JUN-1995; 95US-00467397.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (HYBR-) HYBRIDON INC.
XX Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
DR Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX Claim 1; Page 12; 81pp; English.

CC The present sequence represents a synthetic oligonucleotide HBV92Mb which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a
CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non- contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection

XX Sequence 20 BP; 4 A; 5 C; 6 G; 2 T; 3 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.5;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGCAUGCCATGCC 20
|||||:|||||
Db 1 TAAGGGTCGCAUGCCATGCC 20

RESULT 3
AAT72619
ID AAT72619 standard; DNA; 30 BP.
XX AC AAT72619;
XX 04-SEP-1997 (first entry)
DT Hepatitis B virus RNA antisense oligonucleotide HBV-90Mb.
XX HBV; HBV infection; inhibition; replication; ss.

Key	Location/Qualifiers
FT misc_feature	1..30
FT misc_RNA	1..10
FT modified_base	1
FT modified_base	2
FT modified_base	3
FT modified_base	4
FT modified_base	5

/tag= a
 /note= "Internucleotide linkages are phosphorothioate"
 /tag= b
 /note= "2'-O-Me RNA"
 /tag= c
 /mod_base= OTHER
 /note= "2'-O-methyladenosine"
 /tag= d
 /mod_base= gm
 /tag= e
 /mod_base= OTHER
 /note= "2'-O-methyladenosine"
 /tag= f
 /mod_base= gm
 /tag= g

FT /mod_base= OTHER
 FT modified_base
 FT /note= "2'-O-methyladenosine"
 FT 6
 FT /tag= h
 FT /mod_base= um
 FT modified_base
 FT 7
 FT /tag= i
 FT /mod_base= gm
 FT modified_base
 FT 8
 FT /tag= j
 FT /mod_base= OTHER
 FT modified_base
 FT 9
 FT /note= "2'-O-methyladenosine"
 FT /tag= k
 FT /mod_base= um
 FT modified_base
 FT 10
 FT /tag= l
 FT /mod_base= um
 FT misc_RNA
 FT 21..:30
 FT /tag= m
 FT /note= "2'-OMe RNA"
 FT modified_base
 FT 21
 FT /tag= n
 FT /mod_base= um
 FT modified_base
 FT 22
 FT /tag= o
 FT /mod_base= gm
 FT modified_base
 FT 23
 FT /tag= p
 FT /mod_base= um
 FT modified_base
 FT 24
 FT /tag= q
 FT /mod_base= cm
 FT modified_base
 FT 25
 FT /tag= r
 FT /mod_base= cm
 FT modified_base
 FT 26
 FT /tag= s
 FT /mod_base= OTHER
 FT modified_base
 FT 27
 FT /note= "2'-O-methyladenosine"
 FT /tag= t
 FT /mod_base= um
 FT modified_base
 FT 28
 FT /tag= u
 FT /mod_base= gm
 FT modified_base
 FT 29
 FT /tag= v
 FT /mod_base= cm
 FT modified_base
 FT 30
 FT /tag= w
 FT /mod_base= cm
 FT XX
 FT WO9639502-A1.
 FT PN
 FT XX
 FT 12-DEC-1996.
 FT PD
 FT XX
 FT 04-JUN-1996; 96WO-EP002432.
 FT PF
 FT XX
 FT 06-JUN-1995; 95US-00467397.
 FT PR
 FT XX
 FT (HOFF) HOFFMANN LA ROCHE & CO AG F.
 FT PA
 FT (HYBR-) HYBRIDON INC.
 FT XX
 FT Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
 FT PI
 FT Roberts NA, Roberts PC, Slade A;
 FT XX
 FT WPI; 1997-043124/04.
 FT DR
 FT XX
 FT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 FT used in the detection and treatment of HBV infection.
 FT PT
 FT XX
 FT Claim 5; Page 15; 81pp; English.
 FT PS

XX The present sequence represents a synthetic oligonucleotide HBV-90Mb
 CC which contains a sequence which is complementary to at least two non-
 CC contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
 CC antisense oligonucleotide may be used to detect the presence of HBV in a
 CC sample. The antisense oligonucleotide, and oligonucleotides complementary
 CC to a portion of the HBV RNA, may be used for inhibiting HBV replication
 CC in a cell or for the treatment of HBV infection
 XX
 SQ Sequence 30 BP; 8 A; 5 C; 9 G; 2 T; 6 U; 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 30;
 Best Local Similarity 90.0%; Pred. No. 2.6;
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAAGGGUCCGAUGUCCATGCC 20
 Db 11 TAAGGGTCGAUGUCCATGCC 30
 RESULT 4
 AAT72620
 ID AAT72620 standard; DNA; 30 BP.
 XX AC AAT72620;
 XX 04-SEP-1997 (first entry)
 DT Hepatitis B virus RNA antisense oligonucleotide HBV-91b.
 DE HBV; HBV infection; inhibition; replication; ss.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH misc_feature 1..30
 FT /*tag= a
 FT /note= "Internucleotide linkages are phosphorothioate"
 FT XX
 PN WO9639502-A1.
 XX 12-DEC-1996.
 PD
 PF 04-JUN-1996; 96WO-EP002432.
 XX
 PR 06-JUN-1995; 95US-00467397.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA (HYBR-) HYBRIDON INC.
 XX
 PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
 PI Roberts NA, Roberts PC, Slade A;
 XX
 DR WPI; 1997-043124/04.
 XX
 PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 PT used in the detection and treatment of HBV infection.
 XX
 PS Claim 5; Page 15; 81pp; English.
 CC The present sequence represents a synthetic oligonucleotide HBV-91b which
 CC contains a sequence which is complementary to at least two non-contiguous
 CC regions of a hepatitis B virus (HBV) nucleic acid. The antisense
 CC oligonucleotide may be used to detect the presence of HBV in a sample.
 CC The antisense oligonucleotide, and oligonucleotides complementary to a
 CC portion of the HBV RNA, may be used for inhibiting HBV replication in a
 CC cell or for the treatment of HBV infection
 XX
 SQ Sequence 30 BP; 7 A; 6 C; 11 G; 6 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 30;
 Best Local Similarity 85.0%; Pred. No. 2.6;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 PS

QY 1 TAAGGGUCGAUGUCCATGCC 20
 |||||:||||:|||||
 Db 11 TAAGGGTCGATGTCATGCC 30

RESULT 5

AAT72618

ID AAT72618 standard; DNA; 30 BP.

XX AC AAT72618;

XX DT 04-SEP-1997 (first entry)

XX DE Hepatitis B virus RNA antisense oligonucleotide HBV-90b.
 XX KB HBV; HBV infection; inhibition; replication; ss.
 XX OS Synthetic.

PH Key Location/Qualifiers

FT misc_feature

FT 1..30

FT /note= "Internucleotide linkages are phosphorothioate".

FT /tag= a

FT /note= "Internucleotide linkages are phosphorothioate".

PN WO9639502-A1.

XX 12-DEC-1996.

XX 04-JUN-1996; 96WO-EP002432.

XX 06-JUN-1995; 95US-00467397.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

PA (HYBR-) HYBRIDON INC.

XX Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;

PI Roberts NA, Roberts PC, Slade A;

XX WPI; 1997-043124/04.

XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 used in the detection and treatment of HBV infection.
 PS Claim 5; Page 15; 81pp; English.

CC The present sequence represents a synthetic oligonucleotide HBV-90b which
 contains a sequence which is complementary to at least two non-contiguous
 regions of a hepatitis B virus (HBV) nucleic acid. The antisense
 oligonucleotide may be used to detect the presence of HBV in a sample.
 CC The antisense oligonucleotide, and oligonucleotides complementary to a
 portion of the HBV RNA, may be used for inhibiting HBV replication in a
 cell or for the treatment of HBV infection

XX Sequence 30 BP; 8 A; 5 C; 9 G; 8 T; 0 U; 0 Other;

Query Match

Best Local Similarity 100.0%; Score 20; DB 2; Length 30;

Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCGAUGUCCATGCC 20

|||||:||||:|||||

Db 11 TAAGGGTCGATGTCATGCC 30

RESULT 6

AAT72621

ID AAT72621 standard; DNA; 30 BP.

XX AC AAT72621;

XX DT 04-SEP-1997 (first entry)

XX

DE Hepatitis B virus RNA antisense oligonucleotide HBV-91Mb.

XX HBV; HBV infection; inhibition; replication; ss.
 XX OS Synthetic.

PH Key Location/Qualifiers

FT misc_feature

FT 1..30

FT /note= "Internucleotide linkages are phosphorothioate"

FT 21..30

FT /tag= b

FT /note= "2'-OMe RNA"

FT modified_base

FT 21

FT /tag= c

FT /mod_base= um

FT modified_base

FT 22

FT /tag= d

FT /mod_base= gm

FT modified_base

FT 23

FT /tag= e

FT /mod_base= um

FT modified_base

FT 24

FT /tag= f

FT /mod_base= cm

FT modified_base

FT 25

FT /tag= g

FT /mod_base= cm

FT modified_base

FT 26

FT /tag= h

FT /mod_base= OTHER

FT modified_base

FT 27

FT /tag= i

FT /mod_base= um

FT modified_base

FT 28

FT /tag= j

FT /mod_base= gm

FT modified_base

FT 29

FT /tag= k

FT /mod_base= cm

FT modified_base

FT 30

FT /tag= l

FT /mod_base= cm

FT modified_base

XX WO9639502-A1.

XX 12-DEC-1996.

XX 04-JUN-1996; 96WO-EP002432.

XX 06-JUN-1995; 95US-00467397.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

PA (HYBR-) HYBRIDON INC.

XX Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;

PI Roberts NA, Roberts PC, Slade A;

XX WPI; 1997-043124/04.

XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 used in the detection and treatment of HBV infection.
 PS Claim 5; Page 15; 81pp; English.

CC The present sequence represents a synthetic oligonucleotide HBV-91Mb
 which contains a sequence which is complementary to at least two non-
 contiguous regions of a hepatitis B virus (HBV) nucleic acid. The
 antisense oligonucleotide may be used to detect the presence of HBV in a
 sample. The antisense oligonucleotide, and oligonucleotides complementary
 to a portion of the HBV RNA, may be used for inhibiting HBV replication
 in a cell or for the treatment of HBV infection

```

XX SQ Sequence 30 BP; 7 A; 6 C; 11 G; 3 T; 3 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 90.0%; Pred. No. 2.6;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUGCCATGCC 20
DB 11 TAAGGGTCGAUGCCATGCC 30

RESULT 7
ADC64743/c
ID ADC64743 standard; RNA; 31 BP.
XX AC ADC64743;
XX DT 18-DEC-2003 (first entry)
XX DE Hepatitis B virus DNA polymerase related RNA oligonucleotide.
XX KW screening; antiviral; hepatitis B virus; HBV; DNA polymerase; ss.
XX OS Synthetic.
XX OS Hepatitis B virus.
XX PN KR2002007891-A.
XX PD 29-JAN-2002.
XX PF 19-JUL-2000; 2000KR-00041420.
XX PR 19-JUL-2000; 2000KR-00041420.
XX PA (MOGA-) MOGAM BIOTECHNOLOGY INST.
XX PA (VIRO-) VIROGEN CO LTD.
XX PI Ji HJ, Jung SI, Kim YC, Min MG, Ryu WS, Yoon GS;
XX WPI; 2003-309015/30.
XX DR Screening of antiviral agents by protein-priming activity of hepatitis B
XX PT virus DNA polymerase.
XX PS Disclosure; Page 12; 13pp; Korean.
XX CC The present invention describes a method of screening for an antiviral
XX CC agent by the protein-priming activity of hepatitis B virus (HBV) DNA
XX CC polymerase. Also described is developing an antiviral agent with a high
XX CC selectivity to HBV which can be used for high-throughput screening. The
XX CC present sequence represents an RNA oligonucleotide which is used in the
XX CC exemplification of the present invention.
XX SQ Sequence 31 BP; 10 A; 6 C; 8 G; 0 T; 7 U; 0 Other;
Query Match 100.0%; Score 20; DB 10; Length 31;
Best Local Similarity 85.0%; Pred. No. 2.6;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUGCCATGCC 20
DB 22 TAAGGGTCGATGTCATGCC 3

RESULT 8
AAD27422/c
ID AAD27422 standard; DNA; 639 BP.
XX AC AAD27422;
XX XX
XX DT 18-APR-2002 (first entry)
XX XX

DE Hepatitis B virus (HBV) core antigen (HBcAg) encoding DNA #1.
XX Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;
XX prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;
XX hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;
XX tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;
XX dengue fever; yellow fever; malaria; whooping cough; salmonellosis;
XX food poisoning; meningitis; gonorrhea; antiviral; antibacterial;
XX antiprotozoal; ds.
XX OS Hepatitis B virus.
XX PH Key Location/Qualifiers
XX CDS 1..639
XX FT /*tag= a
XX FT /product= "HBcAg"
XX PN WO200198333-A2.
XX PD 27-DEC-2001.
XX PF 22-JUN-2001; 2001WO-GB002817.
XX PR 22-JUN-2000; 2000GB-00015308.
XX PR 06-OCT-2000; 2000GB-00024544.
XX PA (CELL-) CELLTECH PHARM LTD.
XX PI Page M, Li J, Pumpens P;
XX WPI; 2002-098223/13.
XX DR P-PSDB; AAE17018.
XX PT New proteins comprising a modified hepatitis B core antigen, useful as a
XX PT vaccine in prophylactic or therapeutic vaccination of the human or animal
XX PT body, particularly against hepatitis B virus infection.
XX PS Disclosure; Page 38-39; 40pp; English.
XX CC The invention relates to modified proteins comprising hepatitis B virus
XX CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine
XX CC repeats has been deleted and the protein comprising the C-terminal
XX CC cysteine of HBcAg. The deleted region may be replaced by an epitope from
XX CC a protein other than HBcAg, in which case the HBcAg acts as a carrier to
XX CC present the epitope to the immune system. This chimeric protein or its
XX CC nucleic acid is useful as a vaccine or in a method of prophylactic or
XX CC therapeutic vaccination of the human or animal body, particularly against
XX CC HBV. The nucleic acid encoding the protein may be used in gene therapy or
XX CC DNA vaccination protocols. The chimeric protein or its nucleic acid may
XX CC also be used as the basis of a prophylactic vaccine against a range of
XX CC diseases, e.g. HBV, hepatitis A virus (HAV), hepatitis C virus (HCV),
XX CC influenza, foot-and-mouth disease, polio, herpes, rabies, acquired
XX CC immunodeficiency syndrome (AIDS), dengue fever, yellow fever, malaria,
XX CC tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning,
XX CC diarrhoea, meningitis or gonorrhea. The present sequence is a DNA
XX CC encoding Hepatitis B virus core antigen (HBcAg)
XX SQ Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 6; Length 639;
Best Local Similarity 85.0%; Pred. No. 3.3;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUGCCATGCC 20
DB 104 TAAGGGTCGATGTCATGCC 85

RESULT 9
AAD31509/c
ID AAD31509 standard; DNA; 639 BP.
XX AC AAD31509;

```

```

XX 18-JUN-2002 (first entry)
XX Hepatitis B virus core antigen (HBcAg) encoding DNA.
DE Hepatitis B virus core antigen; HBcAg; prophylactic; viral hepatitis;
XX therapeutic; vaccine; acquired immune deficiency syndrome; influenza;
KW polio; herpes; rabies; AIDS; foot-and-mouth disease; ds.
XX Hepatitis B virus.
XX Key Location/Qualifiers
FT CDS 1..639
FT /tag= a
FT /product= "HBc protein"
FT sig_peptide 1..87
FT /tag= b
FT mat_peptide 88..636
FT /tag= c
FT /product= "Mature HBc protein"
XX WO200177158-A1.
XX 18-OCT-2001.
XX 09-APR-2001; 2001WO-GB001607.
XX 07-APR-2000; 2000EP-00107118.
XX (WEDE-) MEDEVA EURO LTD.
XX Gehin A, Gilbert R, Stuart D, Rowlands D;
XX WPI: 2002-239995/29.
XX P-PSDB; AAE19793.
XX Hepatitis B (HB) core antigen fusion proteins, useful as vaccines for the
PT prophylactic or therapeutic treatment of humans or animals against e.g.
PT HB virus, viral hepatitis, hepatitis C virus, influenza, or foot-and-
PT mouth disease.
XX Disclosure; Page 23-24; 27pp; English.
XX The present invention relates to hepatitis B virus (HBV) core antigen
CC (HBcAg) fusion proteins and polynucleotides encoding such proteins.
CC Sequences of the invention are useful in methods of prophylactic or
CC therapeutic vaccination or to manufacture medicaments for prophylactic or
CC therapeutic vaccination of the human or animal body against HBV, e.g.
CC against viral hepatitis. They are also useful as a prophylactic vaccine
CC against e.g. hepatitis C virus, influenza, polio, herpes, rabies,
CC acquired immune deficiency syndrome (AIDS) or foot-and-mouth disease. The
CC present sequence is a DNA encoding hepatitis B virus core antigen (HBcAg)
XX Sequence 639 BP; 147 A; 161 C; 141 G; 190 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 20; DB 6; Length 639;
Best Local Similarity 85.0%; Pred. No. 3.3;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAGGGUCGAUGUCCATGCC 20
DB 104 TAAGGGTCGATGTCATGCC 85
RESULT 10
AA71734/c
ID AAA71734 standard; cDNA; 663 BP.
XX AAA71734;
XX 06-AUG-2003 (revised)
DT 08-JAN-2001 (first entry)
XX

```

```

DE HBV fusion protein comprising LHB and RGD encoding cDNA.
XX Fusion protein; protein coat; virus-specific packaging signal; psi;
XX virus protein; cell permeability; cell-specific binding site; LHB;
KW large surface protein; core antigen; gene therapy; ss.
XX Hepatitis B virus.
OS Synthetic.
XX WO2000046376-A2.
XX 10-AUG-2000.
XX 04-FEB-2000; 2000WO-DE000363.
XX 05-FEB-1999; 99DE-01004800.
XX (HILD/) HILDT E.
XX Hildt E, Hofschneider P;
XX WPI: 2000-514959/46.
XX P-PSDB; AAB10596.
XX Particle for cell-specific gene delivery, useful in gene therapy,
PT comprises nucleic acid in protein coat that includes a fusion protein of
PT viral protein, permeability peptide and cell-binding site.
XX Claim 16a; Fig 1; 34pp; German.
XX This invention describes a novel particle (A), comprising a protein coat
CC with a fusion protein (FP), and, inside the coat, a nucleic acid (I)
CC including the sequence for a virus-specific packaging signal (psi) and a
CC structural gene. FP contains a virus protein (vp), a peptide (P) that
CC mediates cell permeability and a heterologous cell-specific binding site
CC (RGD). The invention also describes (I) producing (A) in which FP
CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and
CC (P) and RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),
CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector
CC containing the DNA of (d). The products of the invention are used in gene
CC therapy of cells and tissues, in vivo or ex vivo. This sequence encodes a
CC fusion protein which is described in the method of the invention.
XX (Updated on 06-AUG-2003 to correct OS field.)
XX SQ Sequence 663 BP; 154 A; 169 C; 152 G; 188 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 3; Length 663;
Best Local Similarity 85.0%; Pred. No. 3.3;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAGGGUCGAUGUCCATGCC 20
DB 107 TAAGGGTCGATGTCATGCC 88
RESULT 11
AD007220/c
ID ADO07220 standard; DNA; 669 BP.
XX ADO07220;
XX 15-JUL-2004 (first entry)
XX Hepatitis B virus core antigen DNA.
XX HBcAg; immunomodulator; vaccine; gene; ss.
XX Hepatitis B virus.
XX Key Location/Qualifiers
FT CDS 10..669
FT /tag= a
FT /product= "HBcAg"

```

```

FT  /partial
FT  /note= "No start codon"
XX  WO2004035007-A2.
PN  29-APR-2004.
XX  17-OCT-2003; 2003WO-US033178.
XX  17-OCT-2002; 2002US-0419279P.
XX  (ORAG-) ORAGEN CORP.
XX  Michaels F;
PI  WPI; 2004-348329/32.
DR  P-PSDB; ADO07221.
XX  Modulating a systemic immune response to a peptide in a mammal comprises
PT  transmutosally administering a macromolecular aggregate of the peptide.
XX  Disclosure; SEQ ID NO 1; 81pp; English.
XX  The present sequence is the DNA sequence of the hepatitis B virus core
CC  antigen (HBcAg) gene from HBV serotype ayw. A peptide comprising a HBV
CC  protein can be used in claimed methods of the invention for modulating an
CC  immune response in a mammal. A method of inducing a systemic immune
CC  response to a peptide in a mammal comprises transmutosally administering
CC  to the mammal a macromolecular aggregate of the peptide. The
CC  macromolecular aggregate comprises at least 10 peptide subunits, may have
CC  a molecular weight of over 1,000 kDa, and is preferably at least 5 nm in
CC  diameter. It is resistant to digestive degradation, being stabilised in
CC  aggregate form by chemical treatment and/or by recombinant protein
CC  engineering of the peptide. The peptide preferably comprises a HBV
CC  protein selected from HBV surface protein, nucleocapsid protein or
CC  envelope protein. Transmutosomal administration to a mammal of a
CC  macromolecular aggregate of a HBV surface protein engenders a systemic
CC  immune response in the mammal. A method of suppressing an immune response
CC  in a mammal involves transmutosally administering a monomolecular peptide
CC  that is resistant to digestive degradation and which may be stabilised by
CC  chemical treatment or protein engineering, and which may be derived from
CC  a HBV protein. A monomolecular peptide is useful for the induction of
CC  oral tolerance when induction of systemic immunity is undesirable, e.g.
CC  in cases of chronic infections.
XX  SQ  Sequence 669 BP; 155 A; 170 C; 148 G; 196 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 669;
Best Local Similarity 85.0%; Pred. NO. 3.3;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY  1 TAAGGGUCCGAUGCCATGCC 20
DB  |||||:||||:|||||
134 TAAGGTCGATGTCATGCC 115

RESULT 12
AAV82691/C
ID  AAV82691 standard; DNA; 1334 BP.
XX  AAV82691;
AC  AAV82691;
XX  16-FEB-1999 (first entry)
DT  |||||:||||:|||||
DE  134 TAAGGTCGATGTCATGCC 115
XX  Fulminant hepatitis B virus genotype D variant FHBV12 sequence.
XX  Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW  HBV-related disease; ss.
XX  Hepatitis B virus.
OS  Hepatitis B virus.
XX  WO9845421-A2.
XX  PI  Carman B;
XX  WPI; 1999-009329/01.

Query Match 100.0%; Score 20; DB 2; Length 1334;
Best Local Similarity 85.0%; Pred. NO. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY  1 TAAGGGUCCGAUGCCATGCC 20
DB  |||||:||||:|||||
806 TAAGGTCGATGTCATGCC 787

RESULT 13
AAV82688/C
ID  AAV82688 standard; DNA; 1395 BP.
XX  AAV82688;
AC  AAV82688;
XX  16-FEB-1999 (first entry)
DT  |||||:||||:|||||
DE  Fulminant hepatitis B virus genotype D variant FHBV5 sequence.
XX  Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
KW  HBV-related disease; ss.
XX  Hepatitis B virus.
OS  Hepatitis B virus.
XX  WO9845421-A2.
XX  PN  Carman B;
XX  WPI; 1999-009329/01.

```

```

XX New hepatitis B virus nucleic acid with combination of specific mutations
PT - useful for, e.g. detection of binding interactions between host or
PT viral proteins and HBV nucleic.
XX
XX Disclosure; Fig 5; 85pp; English.
XX
XX The present sequence represents part of the genome of a fulminant
CC Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC a mutation (i.e. alteration from the normal nucleotide in any of the
CC genotypes A to F) in at least two of the enhancer II/ core upstream regulatory
CC regulatory element region, the enhancer II/ core upstream regulatory
CC sequence/ basal core promoter region, or a mutation which leads to an X-
CC peptide amino acid change to Cys or Met. The HBV variants of the
CC invention are used to detect binding interactions between host or viral
CC proteins and HBV nucleic acid. Probes that hybridise to any of the
CC specified mutated regions are used to detect HBV-related disease,
CC especially fulminant infection, but also severe chronic infection or
CC serologically unusual forms of disease. Combinations of the specified
CC mutations are associated with fulminant infections, probably because they
CC reduce the ability to bind inhibitory proteins in the host cell
XX
XX Sequence 1395 BP; 277 A; 387 C; 331 G; 398 T; 0 U; 2 Other;
SQ
    Query Match      100.0%; Score 20; DB 2; Length 1395;
    Best Local Similarity 85.0%; Pred. No. 3.5;
    Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
    1 TAAAGGUCGAUGUCCATGCC 20
    |||||:||||:|||||
    917 TAAAGGTCGATGTCATGCC 898

XX
XX RESULT 14
XX AAV82687/c
XX ID AAV82687 standard; DNA; 1400 BP.
XX AC AAV82687;
XX DT 16-FEB-1999 (first entry)
XX DE Fulminant hepatitis B virus genotype D variant FHBV4 sequence.
XX KW Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
XX KW HBV-related disease; ss.
XX OS Hepatitis B virus.
XX PN WO9845421-A2.
XX PD 15-OCT-1998.
XX PF 08-APR-1998; 98WO-EP002048.
XX PR 09-APR-1997; 97GB-00007221.
XX PA (UNIU ) UNIV GLASGOW.
XX PI Carman B;
XX DR WPI; 1999-009329/01.
XX PT New hepatitis B virus nucleic acid with combination of specific mutations
XX PT - useful for, e.g. detection of binding interactions between host or
XX PT viral proteins and HBV nucleic.
XX PS Disclosure; Fig 5; 85pp; English.
XX
XX The present sequence represents part of the genome of a fulminant
CC Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC a mutation (i.e. alteration from the normal nucleotide in any of the
CC genotypes A to F) in at least two of the enhancer II/ core upstream regulatory
CC regulatory element region, the enhancer II/ core upstream regulatory
CC sequence/ basal core promoter region, or a mutation which leads to an X-
CC peptide amino acid change to Cys or Met. The HBV variants of the
CC invention are used to detect binding interactions between host or viral
CC proteins and HBV nucleic acid. Probes that hybridise to any of the
CC specified mutated regions are used to detect HBV-related disease,
CC especially fulminant infection, but also severe chronic infection or
CC serologically unusual forms of disease. Combinations of the specified
CC mutations are associated with fulminant infections, probably because they
CC reduce the ability to bind inhibitory proteins in the host cell
XX
XX Sequence 1395 BP; 277 A; 387 C; 331 G; 398 T; 0 U; 2 Other;
SQ
    Query Match      100.0%; Score 20; DB 2; Length 1395;
    Best Local Similarity 85.0%; Pred. No. 3.5;
    Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
    1 TAAAGGUCGAUGUCCATGCC 20
    |||||:||||:|||||
    917 TAAAGGTCGATGTCATGCC 898

XX
XX RESULT 15
XX AAV82692/c
XX ID AAV82692 standard; DNA; 1445 BP.
XX AC AAV82692;
XX DT 16-FEB-1999 (first entry)
XX DE Fulminant hepatitis B virus genotype D variant FHBV13 sequence.
XX KW Fulminant hepatitis B virus; variant; FHBV; HBV; binding interaction;
XX KW HBV-related disease; ss.
XX OS Hepatitis B virus.
XX PN WO9845421-A2.
XX PD 15-OCT-1998.
XX PF 08-APR-1998; 98WO-EP002048.
XX PR 09-APR-1997; 97GB-00007221.
XX PA (UNIU ) UNIV GLASGOW.
XX PI Carman B;
XX DR WPI; 1999-009329/01.
XX PT New hepatitis B virus nucleic acid with combination of specific mutations
XX PT - useful for, e.g. detection of binding interactions between host or
XX PT viral proteins and HBV nucleic.
XX PS Disclosure; Fig 5; 85pp; English.
XX
XX The present sequence represents part of the genome of a fulminant
CC Hepatitis B virus (FHBV) genotype D variant, nucleotides 1000 to 2500.
CC The specification describes Hepatitis B virus (HBV) nucleic acid that has
CC a mutation (i.e. alteration from the normal nucleotide in any of the
CC genotypes A to F) in at least two of the enhancer II/ core upstream regulatory
CC regulatory element region, the enhancer II/ core upstream regulatory
CC sequence/ basal core promoter region, or a mutation which leads to an X-
CC peptide amino acid change to Cys or Met. The HBV variants of the
CC invention are used to detect binding interactions between host or viral
CC proteins and HBV nucleic acid. Probes that hybridise to any of the
CC specified mutated regions are used to detect HBV-related disease,
CC especially fulminant infection, but also severe chronic infection or
CC serologically unusual forms of disease. Combinations of the specified
CC mutations are associated with fulminant infections, probably because they
CC reduce the ability to bind inhibitory proteins in the host cell
XX
XX Sequence 1400 BP; 287 A; 388 C; 332 G; 393 T; 0 U; 0 Other;
SQ
    Query Match      100.0%; Score 20; DB 2; Length 1400;
    Best Local Similarity 85.0%; Pred. No. 3.5;
    Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
    1 TAAAGGUCGAUGUCCATGCC 20
    |||||:||||:|||||
    917 TAAAGGTCGATGTCATGCC 898

```

XX
SQ Sequence 1445 BP; 297 A; 406 C; 338 G; 404 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 1445;
Best Local Similarity 85.0%; Pred. No. 3.5;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAGGUGCGAUGUCCATGCC 20
DB 917 TAAGGTCGATGTCATGCC 898

Search completed: March 17, 2005, 06:48:43
Job time : 171.333 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-63
Perfect score: 20
Sequence: 1 taagggucauguccatgcc 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 69479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	95.0	686	4	BI954819
2	17.4	87.0	576	8	CC145068
3	17.4	87.0	763	9	CL248902
C 4	17	85.0	260	7	CN589115
5	17	85.0	451	5	BP643309
C 6	17	85.0	729	4	BJ705222
7	16.8	84.0	413	5	BP569691
8	16.8	84.0	431	9	CL260820
9	16.8	84.0	457	5	BP527907
10	16.8	84.0	509	9	TA230D01P
11	16.8	84.0	733	1	AI635930
C 12	16.8	84.0	773	8	BZ573222
C 13	16.8	84.0	1574	9	AG476230
14	16.4	82.0	272	2	BE922196
C 15	16.4	82.0	306	9	AG200800
16	16.4	82.0	356	9	CC836806
C 17	16.4	82.0	417	7	H70178
18	16.4	82.0	499	5	BQ116265
C 19	16.4	82.0	554	8	AQ433745
20	16.4	82.0	612	9	CL194160
21	16.4	82.0	622	5	BQ112668
22	16.4	82.0	633	9	CG461236
23	16.4	82.0	678	9	CG761335
C 24	16.4	82.0	684	9	CL369255

25	16.4	82.0	714	9	AG128210	Pan trogl
26	16.4	82.0	772	8	BH257005	CH230-283
C 27	16.4	82.0	848	7	CO099982	GR_Ra25C
28	16.4	82.0	851	4	BJ571152	BJ571152
29	16.4	82.0	924	9	CG289400	CGXBN55TV
30	16.4	82.0	984	9	CG875596	ZMWBc028
31	16.4	82.0	1010	9	CC822611	CGULP56TV
C 32	16.4	82.0	1046	9	CL055525	CH216-81E
33	16.4	82.0	1669	3	HSMB00819	AL110272 Homo sapi
34	16.4	82.0	1669	3	HSMB00819	AL110272 Homo sapi
35	16	80.0	168	4	BI052244	BI052244 PM2-GN037
36	16	80.0	168	4	BI052309	PM2-GN037
37	16	80.0	532	5	EX433931	EX433931
C 38	16	80.0	666	9	CG106546	CG106546 PUIJ059TB
C 39	16	80.0	698	8	BZ413874	BZ413874 i18d10.9
40	16	80.0	740	9	CG235822	CGMKA72TV
C 41	16	80.0	750	9	CG235810	CGMKA72TH
C 42	16	80.0	810	9	CC716058	CC716058
C 43	16	80.0	836	9	CG257272	CG1CK72TH
44	16	80.0	963	9	CG302315	CG2BM75TV
45	15.8	79.0	68	8	BH631190	BH631190 1007074G0

ALIGNMENTS

RESULT 1
BI954819/c 686 bp mRNA linear EST 19-OCT-2001
LOCUS
DEFINITION
HVSMM0019017f Hordeum vulgare green seedling EST library
HVCNDA0014 (Blumeria infected) Hordeum vulgare subsp. vulgare cDNA
clone HVSMM0019017f, mRNA sequence.

ACCESSION
BI954819
VERSION
BI954819.1 GI:16300646

KEYWORDS
EST.

SOURCE
Hordeum vulgare subsp. vulgare

ORGANISM
Hordeum vulgare subsp. vulgare

REFERENCE
Wing, R., Close, T. J., Klein, H., A., Wise, R., Chin, A., Begum, D.,
Friedrich, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T.,
Simmons, J., Oates, R., and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Blumeria infected Morex (compatible) seedling
cDNA library

AUTHORS
Unpublished (2001)

TITLE
Contact: Wing RA

JOURNAL
Clemson University Genomics Institute

COMMENT
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Seq primer: AATTAACCTCTACTAAGG

Total hg bases = 417

High quality sequence start: 16

High quality sequence stop: 531.

Location/Qualifiers

1. .686

/organism="Hordeum vulgare subsp. vulgare"

/mol_type="mRNA"

/cultivar="Morex"

/sub_species="vulgare"

/db_xref="taxon:112509"

/clone="HVSMM0019017f"

/tissue_type="green seedling leaf"

/lab_host="TJCl21"

/clone_lib="Hordeum vulgare green seedling EST library"

HVCNDA0014 (Blumeria infected)"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; Morex (mla) plants were greenhouse grown in the R

Wise lab at Iowa State University, Ames, IA; 7 day old
Green seedlings were infected with isolate 5874 of
Blumeria graminis f. sp. hordei, and leaves were harvested
24, 48 and 72 hr post-inoculation and snap frozen (Wise).
In the TJ Close lab at the University of California,
Riverside, total RNA was prepared from each sample pool,
equal quantities of all three RNA pools were combined,
poly(A) RNA was purified from the mixture, one primary
unamplified cDNA library was made, and 1 million pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids
(Chin). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
this clone see <http://www.genome.clemson.edu/orders> Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

ORIGIN

Query Match 95.0%; Score 19; DB 4; Length 686;
Best Local Similarity 84.2%; Pred. No. 63;
Matches 16; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAGGGUGCAUGUCCATGCC 20
|||||:||||:|||||
Db 136 AAGGGTCGATGTCATGCC 118

RESULT 2
CC145068
LOCUS
DEFINITION ZMMBBb0002P18.r ZMMBBb Zea mays genomic clone ZMMBBb0002P18 3',
Genomic survey sequence.
ACCESSION CC145068
VERSION CC145068.1 GI:30090261
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 576)
AUTHORS Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.
and Wing, R.

TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0002 row: P column: 18
Seq primer: M13r
Class: BAC ends.

FEATURES
source
1..576
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"

RESULT 4
CN589115/c
LOCUS
DEFINITION TTE00013587 Normalized large Tetrahymena thermophila cDNA, mRNA
sequence.
ACCESSION CN589115

/db_xref="taxon:4577"
/clone="ZMMBBb0002P18"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII; Zea mays L. spp. mays"

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 576;
Best Local Similarity 84.2%; Pred. No. 4.2e+02;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAGGGUGCAUGUCCATGCC 20
|||||:||||:|||||
Db 139 AAGGGTCGATGTCATGCC 157

RESULT 3
CL248902
LOCUS
DEFINITION ZMMBBb0597N17r ZMMBBb (HindIII) Zea mays genomic clone
ZMMBBb0597N17 3', genomic survey sequence.
ACCESSION CL248902
VERSION CL248902.1 GI:41105456
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 763)
AUTHORS Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,
Zohovetz, V., Fuks, G., Yu, Y., Wing, R. and Messing, J.
Sequencing of the maize genome at PGIR (2003c)
Unpublished (2003)
CONTACT: Bharti, A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 405.

FEATURES

source
1..763
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBb0597N17r"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBBb (HindIII)"
/note="Vector: pCUG1; Site_1: HindIII; Site_2: HindIII"

Query Match 87.0%; Score 17.4; DB 9; Length 763;
Best Local Similarity 78.9%; Pred. No. 4.4e+02;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 AAGGGUGCAUGUCCATGCC 20
|||||:||||:|||||
Db 339 AAGGGTCGATGTCATGCC 357

RESULT 4
CN589115/c
LOCUS
DEFINITION TTE00013587 Normalized large Tetrahymena thermophila cDNA, mRNA
sequence.
ACCESSION CN589115

```

VERSION      CN589115.1  GI:47040917
KEYWORDS     EST.
SOURCE       Tetrahymena thermophila
ORGANISM     Tetrahymena thermophila
              Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
              Hymenostomatida; Tetrahymenina; Tetrahymenidae; Tetrahymena.
REFERENCE    1 (bases 1 to 260)
AUTHORS      Garg, J., Pearlman, R.E. and Carlton, J.
TITLE        PEPdbPub (http://amobidida.bcm.umontreal.ca/public/pepdb/agrm.php)
              Tetrahymena thermophila (TIGR)
JOURNAL      Unpublished (2004)
COMMENT      Contact: PEPdb
              Departement de Biochimie, Universite de Montreal
              Email: pepdb-curator@bch.umontreal.ca
              Plate: 1398.

FEATURES     Location/Qualifiers
             1..260
                /organism="Tetrahymena thermophila"
                /mol_type="mRNA"
                /db_xref="taxon:5911"
                /clone_lib="Normalized large"

ORIGIN
Query Match      85.0%; Score 17; DB 7; Length 260;
Best Local Similarity 82.4%; Pred. NO. 6e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGUCCGAGUCCATGCC 20
Db 237 GGGTCGATGTCATGCC 221
|||||:|||||

RESULT 5
BP643309
LOCUS        451 bp mRNA linear EST 27-JUN-2004
DEFINITION  BP643309 RAF19 Arabidopsis thaliana cDNA clone RAF19-61-108 3',
              mRNA sequence.
ACCESSION    BP643309
VERSION      BP643309.1  GI:49294779
KEYWORDS     Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1 (bases 1 to 451)
AUTHORS      Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,
              Nakajima, M., Enju, A., Akiyama, K., Oono, Y., Muramatsu, M.,
              Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y.,
              Arakawa, T., Shibata, K., Shinagawa, A. and Shinozaki, K.
TITLE        Functional annotation of a full-length Arabidopsis cDNA collection
JOURNAL      Science 296 (5565), 141-145 (2002)
MEDLINE      21932900
PUBMED       11910074
COMMENT      Contact: Motoaki Seki
              Plant Functional Genomics Research Group
              RIKEN Genomic Sciences Center
              3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
              Tel: 81-298-36-4359
              Fax: 81-298-36-9060
              Email: mseki@rtc.riken.go.jp
              reversed clone; Please visit our web site
              (http://pfweb.gsc.riken.go.jp/) for further details.
              Location/Qualifiers
             1..451
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
                /clone="RAF19-61-108"
                /tissue_type="mixture of silique and flower"
                /lab_host="DH108"
                /clone_lib="RAF19"
                /notes="Site_1: BamHI; Site_2: SalI; Subtraction Library"

FEATURES     Location/Qualifiers
             1..451
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
                /clone="RAF19-61-108"
                /tissue_type="mixture of silique and flower"
                /lab_host="DH108"
                /clone_lib="RAF19"
                /notes="Site_1: BamHI; Site_2: SalI; Subtraction Library"

ORIGIN
Query Match      85.0%; Score 17; DB 5; Length 451;
Best Local Similarity 82.4%; Pred. NO. 6.5e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 AGGUGCGAGUCCATGCC 19
Db 431 AGGTCGATGTCATGCC 447
|||||:|||||

RESULT 6
BP705222/c
LOCUS        729 bp mRNA linear EST 08-MAR-2004
DEFINITION  BP705222 MF01FFA cDNA Oryzias latipes CDNA clone MF01FFA013a11 5',
              mRNA sequence.
ACCESSION    BP705222
VERSION      BP705222.1  GI:45246102
KEYWORDS     Oryzias latipes (Japanese medaka)
SOURCE       Oryzias latipes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
              Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE    1 (bases 1 to 729)
AUTHORS      Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
TITLE        Medaka EST Project in Takeda's lab
JOURNAL      Unpublished (2001)
COMMENT      Contact: Tadasu Shin-i
              Center For Genetic Resource Information
              National Institute of Genetics
              1111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshin@genes.nig.ac.jp.
              Location/Qualifiers
             1..729
                /organism="Oryzias latipes"
                /mol_type="mRNA"
                /strain="Hd-r"
                /db_xref="taxon:8090"
                /clone="MF01FFA013a11"
                /sex="mixture of female and male"
                /tissue_type="whole embryo"
                /dev_stage="fry stage 40"
                /clone_lib="MF01FFA cDNA"

ORIGIN
Query Match      85.0%; Score 17; DB 4; Length 729;
Best Local Similarity 82.4%; Pred. NO. 7e+02;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 AAGGUGCGAGUCCATG 18
Db 182 AAGGTCGATGTCATG 166
|||||:|||||

RESULT 7
BP569691
LOCUS        413 bp mRNA linear EST 20-JUN-2004
DEFINITION  BP569691 RAF14 Arabidopsis thaliana cDNA clone RAF14-68-M24 3',
              mRNA sequence.
ACCESSION    BP569691
VERSION      BP569691.1  GI:48985457
KEYWORDS     Arabidopsis thaliana (thale cress)
SOURCE       Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1 (bases 1 to 413)
AUTHORS      Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,

```

Nakajima,M., Enju,A., Akiyama,K., Ono,Y., Muramatsu,M.,
 Hayaahizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
 Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
 Functional annotation of a full-length Arabidopsis cDNA collection
 Science 296 (5565), 141-145 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msek@rtc.riken.go.jp

reversed clone; Please visit our web site
 (http://pfweb.gsc.riken.go.jp/) for further details.

FEATURES
 source

1. 431
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL14-68-M24"
 /tissue_type="root"
 /lab_host="DH10B"
 /clone_lib="RAFL14"
 /note="Site_1: BamHI; Site_2: SalI"

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 413;
 Best Local Similarity 75.0%; Pred. No. 8.2e+02;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUCCATGCC 20
 |||||:||||:|||||
 Db 383 TAAGGTGTATGTCATGAC 402

RESULT 8

CL260820 431 bp DNA linear GSS 02-FEB-2004
 ZMWBB0619B24r ZMWBBB (HindIII) Zea mays genomic clone
 ZMWBB0619B24 3', genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

ORGANISM

Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
 Zohovetz,V., Fuku,G., Yu,Y., Wang,R. and Messing,J.

TITLE

JOURNAL

COMMENT

Sequencing of the maize genome at PGIR (2003c)
 Unpublished (2003)
 Contact: Bharti,A.K.
 Dr. Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@waksman.rutgers.edu

Seq primer: SP6

Class: BAC ends

High quality sequence start: 116.

FEATURES

source

1. 431
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="ZMWBB0619B24"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 431;
 Best Local Similarity 80.0%; Pred. No. 8.3e+02;
 Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUCCATGCC 20
 |||||:||||:|||||
 Db 328 TAAGGTGATGCCAGGCC 347

RESULT 9

BP527907 457 bp mRNA linear EST 28-SEP-2004
 BP527907 MAT001 Nicotiana tabacum cDNA clone BY12728, mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Nicotiana tabacum (common tobacco)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
 asterids; lamiids; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 457)

Matsuoka,K., Tashiro,G., Horiguchi,T., Demura,T. and Fukuda,H.

Profiling growth-phase dependent gene expression of tobacco BY-2

cells by comprehensive microarray analysis

Unpublished (2003)

Contact: Ken Matsuoka

Morphogenesis Research Group

RIKEN Plant Science Center

1-7-2 Suehirocho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9575

Fax: 81-45-503-9573

Email: by2@psc.riken.go.jp, URL: http://mrq.psc.riken.go.jp/strc/

The cDNA library was constructed from mRNA isolated from lag (9 h),

seq (72 h) and stationary (7 days) old BY-2 cells.

Seq primer: M13 forward.

Location/Qualifiers

1. 457

/organism="Nicotiana tabacum"

/mol_type="mRNA"

/cultivar="Bright yellow No.2"

/db_xref="taxon:4097"

/clone="BY12728"

/cell_line="BY-2"

/clone_lib="MAT001"

/note="Vector: pGEM-T easy; primer: M13 forward; mRNA

obtained from lag, log and stationary phase cells"

Query Match 84.0%; Score 16.8; DB 5; Length 457;
 Best Local Similarity 75.0%; Pred. No. 8.3e+02;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUCCATGCC 20
 |||||:||||:|||||
 Db 297 TAAAGTTCGATGCCATGCC 316

RESULT 10

TA230D01P 509 bp DNA linear GSS 13-DEC-2000
 LOCUS
 DEFINITION
 T. brucei sheared genomic DNA clone 230d01, forward sequence,
 genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

Trypanosoma brucei

```

ORGANISM Trypanosoma brucei
Bukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 509)
AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
Melville, S.E., Rajandream, M.A. and Barrell, B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
rh@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTAT 10.1) was mechanically sheared
to give a tight size distribution (
4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source             1..509
                        /organism="Trypanosoma brucei"
                        /mol_type="genomic DNA"
                        /strain="TREU927"
                        /db_xref="taxon:5691"
                        /clone="230d01"
ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 509;
Best Local Similarity 75.0%; Pred. No. 8.5e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUGUCCATGCC 20
|||||:|||||
Db 355 TAAGGGTGTGATGCCAGGCC 374

RESULT 11
AI635930
LOCUS t282c11.x1 NCI CGAP Pan1 Homo sapiens cDNA clone IMAGE:2295092 3'
DEFINITION similar to gb:J03490 DIHYDROLIPOAMIDE DEHYDROGENASE PRECURSOR
(HUMAN), mRNA sequence.
ACCESSION AI635930
VERSION AI635930.1 GI:4687260
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 733)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-x@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
found through the NCI-CGAP clone distribution information can be
www-bio.llnl.gov/bbr/image/image.html
Insert Length: 1107 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 401.
Location/Qualifiers
     source             1..733
                        /organism="Homo sapiens"

Trypanosoma brucei
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2295092"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pan1"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"
ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 733;
Best Local Similarity 75.0%; Pred. No. 8.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUGUCCATGCC 20
|||||:|||||
Db 557 TAAGGTCGATGTCATGAC 576

RESULT 12
BZ573222/c
LOCUS msh2_3006.y2 msh Pseudomonas aeruginosa genomic clone msh2_3006,
DEFINITION genomic survey sequence.
ACCESSION BZ573222
VERSION BZ573222.1 GI:27208283
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 773)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES             Location/Qualifiers
     source             1..773
                        /organism="Pseudomonas aeruginosa"
                        /mol_type="genomic DNA"
                        /strain="MSH"
                        /db_xref="taxon:287"
                        /clone="msh2_3006"
                        /clone_lib="msh"
                        /note="Environmental isolate. Whole genomic shotgun
library."
ORIGIN
Query Match      84.0%; Score 16.8; DB 8; Length 773;
Best Local Similarity 75.0%; Pred. No. 9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGUCCGAUGUCCATGCC 20
|||||:|||||
Db 294 TAATGTCGATGTCAGGCC 275

RESULT 13
AG476230/c
LOCUS AG476230
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-369C21.TJ, genomic survey
sequence.

```

ACCESSION AG476230
 VERSION AG476230.1 GI:48183460
 KEYWORDS GSS.
 SOURCE Mus musculus molossinus
 ORGANISM Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
 TITLE BAC end Sequences of Library MSMg01
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1574)
 AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9114, Fax:81-45-503-9170)
 COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@tc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan
 phone: 81-298-36-9189, fax: 81-298-36-9199
 e-mail: abe@tc.riken.jp
 PRIMERS
 Sequencing: TJ
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.

FEATURES source
 Location/Qualifiers
 1..1574
 /organism="Mus musculus molossinus"
 /mol_type="genomic DNA"
 /sub_species="molossinus"
 /db_xref="taxon:57486"
 /clone="MSMg01-369C21.TJ"
 /sex="male"
 /tissue_type="mixture of kidney and spleen"
 /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 9; Length 1574;
 Best Local Similarity 75.8%; Pred.No.1e+03; 2; Indels 0; Gaps 0;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TAAGGGUGCAUGUCCATGCC 20
 |||:|||||
 Db 300 TATGGGTCGATGTCATCCC 281

RESULT 14
 BE922196
 LOCUS
 DEFINITION EST425953 potato leaves and petioles Solanum tuberosum cDNA clone
 CSTR18G4 5' sequence, mRNA sequence.
 ACCESSION BE922196
 VERSION BE922196.1 GI:10448260
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 272)
 AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J., Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
 TITLE Generation of ESTs from potato leaves and petioles
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robin Buell

The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/.

FEATURES source
 Location/Qualifiers
 1..272
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="CSTB18G4"
 /tissue_type="leaflets and petioles"
 /dev_stages="8 weeks old plants"
 /lab_host="SOLR"
 /clone_lib="potato leaves and petioles"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI. Tissue was supplied by Dr. Fry (Cornell University). Leaflets and petioles were isolated from 8 week old greenhouse grown plants. The plants were watered and fertilized freely. The tissue was immediately frozen in liquid nitrogen."

ORIGIN
 Query Match 82.0%; Score 16.4; DB 2; Length 272;
 Best Local Similarity 77.8%; Pred.No.1.2e+03;
 Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 AGGGUGCAUGUCCATGCC 20
 |||:|||||
 Db 29 AGGTTGATGTCATGCC 46

RESULT 15
 AG200800/c
 LOCUS
 DEFINITION Pan troglodytes DNA, clone: RP43-082P16.T7, genomic survey
 sequence.
 ACCESSION AG200800
 VERSION AG200800.1 GI:45232975
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1
 AUTHORS Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J., Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
 TITLE BAC end sequences of Library RP-43
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 306)
 AUTHORS Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J., Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
 TITLE Direct Submission
 JOURNAL Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of Bioscience and Biotechnology (KRIIB), Genome Research Center (GRC); 52, Oun-dong, Yuseong-gu, Daejeon 305-333, Korea (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/, Tel:82-42-866-7181, Fax:82-42-860-4409)
 COMMENT Clones are derived from the chimpanzee BAC library RP-43 This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
 PRIMERS
 Sequencing: T7
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.

FEATURES source
 Location/Qualifiers
 1..306
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"

```
/db_xref="taxon:9598"  
/clone="RP43-082P16.T7"  
/sex="male"  
/cell_type="lymphocytes"  
/clone_lib="RP-43 Chimpanzee Male BAC Library"
```

ORIGIN

Query Match	82.0%;	Score 16.4;	DB 9;	Length 306;
Best Local Similarity	77.8%;	Pred. NO. 1.3e+03;		
Matches 14;	Conservative	3;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	TAAGGGUCGAUCCCATG	18	
		: : :		
Db	35	TAAGGGTAGATGTCCCATG	18	

Search completed: March 17, 2005, 11:07:49
Job time : 1386.27 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-64

Perfect score: 20

Sequence: 1 ttataagggtcgauccau 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6	AR027818
2	20	100.0	21	6	AX137490 Sequence
3	20	100.0	21	6	BD011606
4	20	100.0	27	6	AR027819
5	20	100.0	28	6	AR103926
6	20	100.0	35	6	BD236992 DNA vacci
7	20	100.0	81	6	I92348
8	20	100.0	253	14	AY329529
9	20	100.0	253	14	AY329561
10	20	100.0	253	14	AY329562
11	20	100.0	253	14	AY329568
12	20	100.0	253	14	AY329573
13	20	100.0	253	14	AY329575
14	20	100.0	253	14	AY329581
15	20	100.0	294	14	AF390000
16	20	100.0	333	14	HP8BBD
17	20	100.0	398	14	AB167603
18	20	100.0	398	14	AB167637
19	20	100.0	406	14	AB163815

C 20	100.0	406	14	AB163817	Hepatitis
C 21	100.0	439	14	AY254503	Hepatitis
C 22	100.0	456	14	AY509204	Hepatitis
C 23	100.0	488	14	AY274419	Hepatitis
C 24	100.0	488	14	AY274420	Hepatitis
C 25	100.0	488	14	AY274422	Hepatitis
C 26	100.0	488	14	AY274427	Hepatitis
C 27	100.0	488	14	AY274428	Hepatitis
C 28	100.0	488	14	AY274429	Hepatitis
C 29	100.0	488	14	AY274430	Hepatitis
C 30	100.0	488	14	AY274431	Hepatitis
C 31	100.0	488	14	AY274432	Hepatitis
C 32	100.0	488	14	AY274433	Hepatitis
C 33	100.0	488	14	AY274434	Hepatitis
C 34	100.0	488	14	AY274436	Hepatitis
C 35	100.0	548	14	AY382500	Hepatitis
C 36	100.0	548	14	AY382501	Hepatitis
C 37	100.0	548	14	AY382502	Hepatitis
C 38	100.0	548	14	AY382521	Hepatitis
C 39	100.0	548	14	AY382522	Hepatitis
C 40	100.0	548	14	AY382523	Hepatitis
C 41	100.0	548	14	AY382524	Hepatitis
C 42	100.0	548	14	AY382525	Hepatitis
C 43	100.0	548	14	AY382526	Hepatitis
C 44	100.0	548	14	AY382527	Hepatitis
C 45	100.0	552	6	BD236991	BD236991 DNA vacci

ALIGNMENTS

RESULT 1	AR027818	Sequence 16 from patent US 5856459.	20 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	AR027818					
DEFINITION	Sequence 16 from patent US 5856459.					
ACCESSION	AR027818					
VERSION	AR027818.1	GI:5938638				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 20)					
AUTHORS	Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.					
TITLE	Oligonucleotides specific for Hepatitis B virus					
JOURNAL	Patent: US 5856459-A 16 05-JAN-1999;					
FEATURES	Location/Qualifiers					
source	1..20					
ORIGIN	/organism="unknown"					
	/mol_type="unassigned DNA"					

Query Match	100.0%;	Score 20;	DB 6;	Length 20;
Best Local Similarity	85.0%;	Pred. No. 2;		
Matches	17;	Conservative 3;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	TTATAAGGGTCGAUGUCCAU 20		
Db	1	TTATAAGGGTCGATGCCAT 20		
RESULT 2				
LOCUS	AX137490	Sequence 3 from Patent EP1072271.	21 bp	DNA
DEFINITION	AX137490			linear
ACCESSION	AX137490			
VERSION	AX137490.1	GI:14273684		
KEYWORDS				
SOURCE	Hepatitis B virus			
ORGANISM	Hepatitis B virus			
REFERENCE	1	Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.		
AUTHORS	Rabbani,E., Ilan,Y., Roy-Chowdhury,J. and Engelhardt,D.L.			

TITLE Selective immune down regulation (sidr) mediated transplantation processes
JOURNAL Patent: EP 1072271-A 3 31-JAN-2001;
ENZO THERAPEUTICS, INC. (US)
FEATURES
source
1. .21
/organism="Hepatitis B virus"
/mol_type="unassigned DNA"
/db_xref="taxon:10407"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 21;
Best Local Similarity 85.0%; Pred. No. 2;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATAAGGTCGAUGUCCAU 20
Db 2 TTATAAGGTCGATGCCAT 21
RESULT 3
BD011606
LOCUS 21 bp DNA linear PAT 09-JAN-2004
DEFINITION Biological models showing secondary disease sign and useful in developing remedies, diagnostic products and therapeutic or diagnostic procedure, method with the use of the same and cells, tissues and organs derived therefrom.
ACCESSION BD011606
VERSION 1
KEYWORDS JP 2001078621-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Brown, J.J., Rabbani, I., Donegan, J.J. and Chaudhury, J.R.
TITLE Biological models showing secondary disease sign and useful in developing remedies, diagnostic products and therapeutic or diagnostic procedure, method with the use of the same and cells, tissues and organs derived therefrom
JOURNAL Patent: JP 2001078621-A 3 27-MAR-2001;
ENZO THERAPEUTICS INC
COMMENT OS Hepatitis virus (hepatitis B virus)
PN JP 2001078621-A/3
PD 27-MAR-2001
PF 14-JUL-2000 JP 2000215182
PR 16-JUL-1999 US 09/356293
PI JENNIFER JUNE BROWN, IRAZAR RABBANI, JAMES J DONEGAN, PI JAYANTA ROY CHAUDHURY
PC A01K67/027, A61K45/00, C12N5/00, C12N15/09, C12Q1/68, G01N33/15, PC G01N33/50//
PC C12N1/20, C12N7/00, C12N5/00, C12N15/00
CC C12N1/20, C12N7/00, C12N5/00, C12N15/00
FH Key Location/Qualifiers
FT source 1. .21
FT /organism="Hepatitis virus (hepatitis B virus)".
FEATURES
source
1. .21
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 21;
Best Local Similarity 85.0%; Pred. No. 2;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATAAGGTCGAUGUCCAU 20
Db 2 TTATAAGGTCGATGCCAT 21
RESULT 4
AR027819

LOCUS AR027819 27 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 17 from patent US 5856459.
ACCESSION AR027819
VERSION AR027819.1 GI:5938639
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Frank, B.L., Roberts, P.C., Goodchild, J., Craig, J., Charles. and Mills, J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 17 05-JAN-1999;
FEATURES Location/Qualifiers
1..27
source /organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 27;
Best Local Similarity 85.0%; Pred. No. 2.1;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATAAGGTCGAUGUCCAU 20
Db 8 TTATAAGGTCGATGCCAT 27
RESULT 5
AR103926/c
LOCUS AR103926 28 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6087556.
ACCESSION AR103926
VERSION AR103926.1 GI:12815514
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Feitelson, M. and Siracusa, L.
TITLE Transgenic animals capable of replicating hepatitis viruses and mimicking chronic hepatitis infection in humans
JOURNAL Patent: US 6087556-A 1 11-JUL-2000;
FEATURES Location/Qualifiers
1..28
source /organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 28;
Best Local Similarity 85.0%; Pred. No. 2.1;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATAAGGTCGAUGUCCAU 20
Db 20 TTATAAGGTCGATGCCAT 1
RESULT 6
BD236992/c
LOCUS BD236992 35 bp DNA linear PAT 17-JUL-2003
DEFINITION DNA vaccination to cholesterol ester transfer protein in the treatment of atherosclerosis.
ACCESSION BD236992
VERSION BD236992.1 GI:33046762
KEYWORDS JP 2002516656-A/17.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 35)
AUTHORS Needleman, P. and Glenn, K.
TITLE DNA vaccination to cholesterol ester transfer protein in the treatment of atherosclerosis

INISIM
hepatitis B virus
Hepatitis B virus
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

```

/isolate="D272811E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95924.1"
/db_xref="GI:37625411"
/translation="STTDLEAYFKDCLFKDWEELGEEYELMIFVLGGCRHKLVCAPAP
CNFTSA"
134..217
/codon_start=1
/product="prec/C protein"
/protein_id="AAQ95925.1"
/db_xref="GI:37625412"
/translation="MQLFHLCLIISCSCTVQASKLCLGWL"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 2.9;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
|||||
Db 240 TTATAAGGTCGATGCCAT 221

RESULT 10
AY329562/c
LOCUS
DEFINITION
Hepatitis B virus isolate D273984E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329562
VERSION
AY329562.1 GI:37625413
KEYWORDS
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
15184419
PUBMED
2 (bases 1 to 253)
Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratorio.
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
LOCATION/Qualifiers
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D273984E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95938.1"
/db_xref="GI:37625432"
/translation="STTDLEAYFKDCLFKDWEELGEEELRLMIFVLGGCRHKLVCAPAP
CNFTSA"
134..217
/codon_start=1
/product="prec/C protein"
/protein_id="AAQ95939.1"
/db_xref="GI:37625433"
/translation="MQLFHLCLIISCSCTVQASKLCLGWL"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 2.9;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
|||||
Db 240 TTATAAGGTCGATGCCAT 221

RESULT 12
AY329573/c
LOCUS
DEFINITION
Hepatitis B virus isolate D604917E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329573
VERSION
AY329573.1 GI:37625446
KEYWORDS
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
15184419
PUBMED
2 (bases 1 to 253)
Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratorio
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
LOCATION/Qualifiers
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D273984E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95926.1"
/db_xref="GI:37625414"
/translation="STTDLEAYFKDCLFKDWEELGEEIRLMIFVLGGCRHKLVCAPAP
CNFTSA"
134..217
/codon_start=1
/product="prec/C protein"
/protein_id="AAQ95927.1"
/db_xref="GI:37625415"
/translation="MQLFHLCLIISCSCTVQASKLCLGWL"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 2.9;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
|||||
Db 240 TTATAAGGTCGATGCCAT 221

RESULT 11
AY329568/c
LOCUS
DEFINITION
Hepatitis B virus isolate D29668E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329568
VERSION
AY329568.1 GI:37625431
KEYWORDS
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
15184419
PUBMED
2 (bases 1 to 253)
Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratorio.
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
LOCATION/Qualifiers
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D29668E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95938.1"
/db_xref="GI:37625432"
/translation="STTDLEAYFKDCLFKDWEELGEEELRLMIFVLGGCRHKLVCAPAP
CNFTSA"
134..217
/codon_start=1
/product="prec/C protein"
/protein_id="AAQ95939.1"
/db_xref="GI:37625433"
/translation="MQLFHLCLIISCSCTVQASKLCLGWL"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 2.9;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
|||||
Db 240 TTATAAGGTCGATGCCAT 221

RESULT 13
AY329573/c
LOCUS
DEFINITION
Hepatitis B virus isolate D604917E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329573
VERSION
AY329573.1 GI:37625446
KEYWORDS
SOURCE
Hepatitis B virus
ORGANISM
Hepatitis B virus
VIRUSES; Retroviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
SITNIK,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
15184419
PUBMED
2 (bases 1 to 253)
Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratorio
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
LOCATION/Qualifiers
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D604917E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95927.1"
/db_xref="GI:37625415"
/translation="MQLFHLCLIISCSCTVQASKLCLGWL"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 2.9;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
|||||
Db 240 TTATAAGGTCGATGCCAT 221

```

```

/mol_type="genomic DNA"
/isolate="D611058E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95952.1"
/db_xref="GI:37625453"
/translation="STTDLEAYFKDCLFKDWEELGBETRLMI FVLGGCRHKLVCAPAP
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95953.1"
/db_xref="GI:37625454"
/translation="MQLFHLCLIISCSCTPTQVASKLCLGWL"

ORIGIN

Query Match      100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 2.9;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1  TTATAAGGTCGAUGUCCAU 20
      |||||
Db      240 TTATAAGGTCGATGCCAT 221

RESULT 14
AY329581/c
LOCUS
DEFINITION
Hepatitis B virus isolate D639472E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329581
VERSION
AY329581.1 GI:37625470
KEYWORDS
Hepatitis B virus
SOURCE
Hepatitis B virus
ORGANISM
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
AUTHORS
Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
TITLE
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
15184419
2 (bases 1 to 253)
AUTHORS
Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUN-2003) Research & Development, Laboratorio
Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
Location/Qualifiers
1..253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D639472E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AAQ95964.1"
/db_xref="GI:37625471"
/translation="STTDLEAYFKDCLFKDWEELGBELRLLI FVLGGCRHKLVCAPAP
CNFF TSA"
134..217
/codon_start=1
/product="preC/C protein"
/protein_id="AAQ95965.1"
/db_xref="GI:37625472"
/translation="MQLFHLCLIISCSCTPTQVASKLCLGWL"

ORIGIN

```

```

Query Match      100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 85.0%; Pred. No. 2.9;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTATAAGGTCGAUGUCCAU 20
      |||||
Db      240 TTATAAGGTCGATGCCAT 221

RESULT 15
AF390000/c
LOCUS      294 bp DNA linear VRL 06-MAR-2002
DEFINITION Hepatitis B virus isolate D3 X protein gene, partial cds; and
nonfunctional precore/core protein gene, partial sequence.
ACCESSION AF390000
VERSION AF390000.1 GI:16266099
KEYWORDS
SOURCE
ORGANISM Hepatitis B virus
Hepatitis B virus
Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE 1 (bases 1 to 294)
AUTHORS Castro,L.D., Niel,C. and Gomes,S.A.
TITLE Low frequency of mutations in the core promoter and precore regions
of hepatitis B virus in anti-HBe positive Brazilian carriers
JOURNAL BMC Microbiol. 1 (1), 10 (2001)
PUBMED 11472634
REFERENCE 2 (bases 1 to 294)
AUTHORS De Castro,L., Niel,C. and Gomes,S.A.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-2001) Virology, FIOCRUZ, Av. Brasil 4365, Rio de
Janeiro, RJ 21045-900, Brazil
FEATURES
source
1..294
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D3"
/db_xref="taxon:10407"
<1..119
/codon_start=3
/product="X protein"
/protein_id="AAU16752.1"
/db_xref="GI:16266100"
/translation="FKDWELGDDSLMIYVLGCRHKLVCAAPCNFTTSA"
95..>294
misc_feature
/note="nonfunctional precore/core protein due to mutation"

ORIGIN

Query Match      100.0%; Score 20; DB 14; Length 294;
Best Local Similarity 85.0%; Pred. No. 3;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTATAAGGTCGAUGUCCAU 20
      |||||
Db      201 TTATAAGGTCGATGCCAT 182

Search completed: March 17, 2005, 08:14:18
Job time : 684.733 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-64
Perfect score: 20
Sequence: 1 ttataagggtcgauguccau 20
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	2 AAT72575	Aat72575 Hepatitis
2	20	100.0	20	2 AAT72576	Aat72576 Hepatitis
3	20	100.0	21	2 AAQ38441	Aaq38441 Antisense
4	20	100.0	21	3 AAZ98697	Aaz98697 Human hep
5	20	100.0	21	5 AAF58672	Aaf58672 Hepatitis
6	20	100.0	21	10 ADB68574	Adb68574 NG2 A-L-P
7	20	100.0	21	10 ADB68572	Adb68572 A-L-P con
8	20	100.0	27	2 AAT72577	Aat72577 Hepatitis
9	20	100.0	28	2 AAQ80499	Aaq80499 Primer to
10	20	100.0	28	2 AAV45201	Aav45201 Primer MF
11	20	100.0	28	3 AAAG2585	Aag2585 Transgeni
12	20	100.0	28	12 ADN36070	Adn36070 Probe #15
13	20	100.0	28	12 ADN36068	Adn36068 Probe #14
14	20	100.0	28	12 ADN36071	Adn36071 Probe #15
15	20	100.0	28	12 ADN36073	Adn36073 Probe #15
16	20	100.0	28	12 ADN36066	Adn36066 Probe #14
17	20	100.0	28	12 ADN36067	Adn36067 Probe #14
18	20	100.0	28	12 ADN36072	Adn36072 Probe #15
19	20	100.0	28	12 ADN36065	Adn36065 Probe #14
20	20	100.0	28	12 ADN36069	Adn36069 Probe #15

C 21	20	100.0	31	6 ABA96791	Abas96791 Hepatitis
C 22	20	100.0	31	10 ADC64743	Adc64743 Hepatitis
C 23	20	100.0	35	2 AAX36590	Aax36590 PCR prime
C 24	20	100.0	35	8 ABX95880	Abx95880 PCR prime
C 25	20	100.0	35	10 ACD07807	Ac07807 Hepatitis
C 26	20	100.0	36	10 ADJ94539	Adj94539 HBV genom
C 27	20	100.0	39	13 ADR89273	Adr89273 Lab-on-ch
C 28	20	100.0	39	13 ADR89266	Adr89266 Lab-on-ch
C 29	20	100.0	53	12 ADN36055	Adn36055 Probe #13
C 30	20	100.0	504	11 ADM41005	Adm41005 Hbc relat
C 31	20	100.0	504	11 ADM41004	Adm41004 Hbc relat
C 32	20	100.0	513	6 ABK67524	Abk67524 DNA encod
C 33	20	100.0	513	6 ABK67525	Abk67525 DNA encod
C 34	20	100.0	516	6 ABK67527	Abk67527 DNA encod
C 35	20	100.0	519	6 ABK67526	Abk67526 DNA encod
C 36	20	100.0	534	11 ADM40998	Adm40998 Hbc relat
C 37	20	100.0	534	11 ADM40999	Adm40999 Hbc relat
C 38	20	100.0	534	11 ADM41007	Adm41007 Hbc relat
C 39	20	100.0	534	11 ADM41008	Adm41008 Hbc relat
C 40	20	100.0	540	11 ADM41010	Adm41010 Hbc relat
C 41	20	100.0	540	11 ADM41011	Adm41011 Hbc relat
C 42	20	100.0	549	6 ABK44278	Abk44278 DNA encod
C 43	20	100.0	549	6 ABK67533	Abk67533 Immunogen
C 44	20	100.0	549	10 ADE10968	Adel0968 Human Hep
C 45	20	100.0	549	10 ADG47010	Adg47010 Hepatitis

ALIGNMENTS

RESULT 1
AAT72575
ID AAT72575 standard; DNA; 20 BP.
XX
AC AAT72575;
XX
DT 04-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV101b.
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..20
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"

XX
PN WO9639502-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP002432.
XX
PR 06-JUN-1995; 95US-00467397.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (HYBR-) HYBRIDON INC.
XX
PI Craig CV, Frank BL, Goodchild J, Jupp R, Kiluskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
XX
WPI; 1997-043124/04.
XX
PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
PS Claim 1; Page 12; 81pp; English.
XX
CC The present sequence represents a synthetic oligonucleotide HBV101b which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a

Query Match 100.0%; Score 20; DB 2; Length 21;
 Best Local Similarity 85.0%; Pred. No. 1.2;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
 Db 2 TTATAAGGTCGATGCCAT 21

RESULT 4
 AA298697
 ID AA298697 standard; DNA; 21 BP.
 XX
 AC AA298697;
 XX
 DT 06-AUG-2003 (revised)
 DT 20-JUN-2000 (first entry)
 XX
 DE Human hepatitis B virus antisense oligonucleotide sequence.
 XX
 KW Cytostatic; virucide; hepatotropic; anti-inflammatory; antisense;
 KW hepatitis B virus; oncoprotein expression inhibitor; ss;
 KW asialoglycoprotein receptor.
 XX
 OS Hepatitis B virus.
 XX
 PN US6030954-A.
 XX
 PD 29-FEB-2000.
 XX
 PF 02-JUN-1995; 95US-00459633.
 XX
 PR 05-SEP-1991; 91US-00755083.
 PR 04-NOV-1991; 91US-00788119.
 PR 03-APR-1992; 92US-00864003.
 PR 04-SEP-1992; 92US-00941366.
 XX
 PA (UYCO-) UNIV CONNECTICUT.
 XX
 PI Wu CH, Wu GY;
 XX
 DR WPI; 2000-223192/19.
 XX
 PT Antisense oligonucleotides targeted to asialoglycoprotein receptor-
 bearing cells useful for inhibiting viral and oncoprotein RNA expression.
 XX
 PS Example 1; Col 5; 12pp; English.
 XX
 CC This sequence represents a human hepatitis B antisense oligonucleotide
 which can be used as a component of the soluble molecular complex of the
 invention. The invention relates to a soluble molecular complex
 comprising a single stranded antisense oligonucleotide which hybridises
 to an RNA in a target cell. The antisense oligonucleotide is complexed
 with a carrier comprised of a ligand for the asialoglycoprotein receptor
 and a polycation. The molecular complex has cytostatic, virucide,
 hepatotropic and anti-inflammatory activity. The complex works through
 cell specific antisense inhibition of RNA expression. The molecular
 complex is used for inhibiting oncogene and viral (especially hepatitis)
 RNA expression in asialoglycoprotein receptor-bearing cells. (Updated on
 06-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 21 BP; 5 A; 3 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 21;
 Best Local Similarity 85.0%; Pred. No. 1.2;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
 Db 2 TTATAAGGTCGATGCCAT 21

RESULT 5

AA258672
 ID AA258672 standard; DNA; 21 BP.
 XX
 AC AA258672;
 XX
 DT 27-APR-2001 (first entry)
 XX
 DE Hepatitis B virus HB probe.
 XX
 KW Hepatitis B virus; HBV; antiinflammatory; immunosuppressive;
 KW hepatotropic; virucide; vaccine; Tupaia belangeri;
 KW immune-related disorder; transplantation rejection;
 KW selective immune down regulation; SDR; probe; ss.
 XX
 OS Hepatitis B virus.
 XX
 PN EP1072271-A2.
 XX
 PD 31-JAN-2001.
 XX
 PF 17-JUL-2000; 2000EP-00115423.
 XX
 PR 16-JUL-1999; 99US-00356294.
 XX
 PA (ENZO-) ENZO THERAPEUTICS INC.
 XX
 PI Rabbani E, Ilan Y, Roy-Chowdhury J, Engelhardt DL;
 XX
 DR WPI; 2001-170934/18.
 XX
 PT Native or non-native antigens, used for establishing selective immune
 down regulation, for transplantation, for treating or preventing
 undesirable immune reactions of vaccination, and for treating immune
 disorders.
 XX
 PS Example 7; Page 18; 47pp; English.
 XX
 CC The present sequence is a probe which was used in an example
 demonstrating disease symptoms induced in a small animal model, Tupaia
 belangeri, after infection by Hepatitis B virus (HBV). This example is
 provided in a specification relating to the use of native or non-native
 antigen or antigens, or their immunological equivalent, for preparing a
 pharmaceutical composition for use in transplantation processes, for
 treating or preventing undesirable immunological consequences of
 vaccination or immunisation in a subject, or for treating immune-related
 disorders. The invention provides unique selective immune down regulation
 (SIDR) applications in transplantation processes. They may be used for
 preventing or treating graft versus host rejection and for treating
 Crohn's disease, primary sclerosing cholangitis disease, primary biliary
 cirrhosis disease, primary Celliac's disease, primary autoimmune chronic
 active hepatitis, chronic liver rejection disease, immune-mediated liver
 fibrosis disease, immune-mediated vascular disorder, and immune-mediated
 muscle disorders affecting smooth muscle, striated muscle and blood
 vessel muscle
 XX
 SQ Sequence 21 BP; 5 A; 3 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 5; Length 21;
 Best Local Similarity 85.0%; Pred. No. 1.2;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
 Db 2 TTATAAGGTCGATGCCAT 21

RESULT 6
 ADB68574
 ID ADB68574 standard; DNA; 21 BP.
 XX
 AC ADB68574;
 XX
 DT 04-DEC-2003 (first entry)

XX NG2 A-L-P conjugate DNA component used to target HBV c-gene.
 XX homogeneous A-L-P conjugate; hepatic; chronic viral hepatitis; cirrhosis;
 KW malaria; viral infection; protozoan; cancer; hepatocellular carcinoma;
 KW HCC; ss; HBV; c-gene; core.
 XX Hepatitis B virus.
 XX OS
 XX Location/Qualifiers
 FH Key
 FT modified_base 1..21
 FT /tag= b
 FT /mod_base= OTHER
 FT /note= "OTHER = phosphorothioate backbone"
 FT modified_base 1
 FT /tag= a
 FT /mod_base= OTHER
 FT /note= "OTHER = Optionally linked to YEE(ahGalNAc)3-SMCC
 FT and various chemical groups as shown in figures"
 FT modified_base 21
 FT /tag= c
 FT /mod_base= OTHER
 FT /note= "OTHER = Optionally linked to chemical group as
 FT shown in figure 5"
 FT XX
 XX WO2003067209-A2.
 XX 14-AUG-2003.
 XX 21-JUN-2002; 2002WO-US019908.
 XX 22-JUN-2001; 2001US-00888164.
 XX (CELL-) CELL WORKS INC.
 XX (UYJO) UNIV JOHNS HOPKINS.
 XX Ts'o POP, Duff R, Zhou Y, Deamond S, Roby C;
 XX WPI; 2003-697456/66.
 XX New homogeneous prodrug conjugate containing hepatic ligand for delivery
 PT of pathogen-specific oligomer useful for treating liver infections or
 PT cancer.
 XX Claim 7; Page 83; 107pp; English.
 XX The invention relates to a novel homogeneous conjugate comprising a
 CC hepatic ligand, bifunctional linker and biologically stable oligomer that
 CC binds to a sequence in a hepatic virus or pathogen and is released from
 CC the conjugate by hydrolysis or reduction. The conjugate of the invention
 CC may be useful during the treatment of liver diseases including chronic
 CC viral hepatitis, cirrhosis, malaria, viral or protozoan infection and
 CC cancer, such as hepatocellular carcinoma (HCC). The current sequence is
 CC that of the NG2 A-L-P conjugate DNA component of the invention which was
 CC used to target the Hepatitis B virus (HBV) c (core)-gene.
 XX Sequence 21 BP; 5 A; 3 C; 5 G; 8 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 20; DB 10; Length 21;
 Best Local Similarity 85.0%; Pred. No. 1.2;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATAAGGGTCGAUGUCCAU 20
 |||||
 2 TTATAAGGGTCGATGCCAT 21
 Db
 RESULT 7
 ID ADB68572
 AC ADB68572 standard; RNA; 21 BP.
 XX ADB68572;
 XX

DT 04-DEC-2003 (first entry)
 XX A-L-P conjugate-related RNA oligomer 3.
 XX homogeneous A-L-P conjugate; hepatic; chronic viral hepatitis; cirrhosis;
 KW malaria; viral infection; protozoan; cancer; hepatocellular carcinoma;
 KW HCC; ss.
 XX Unidentified.
 XX OS
 XX Location/Qualifiers
 FH Key
 FT modified_base 1..21
 FT /tag= a
 FT /mod_base= OTHER
 FT /note= "OTHER = 2'-O-methyl ribose alternating methyl-
 FT phosphate-phosphodiester backbone"
 FT XX
 XX WO2003067209-A2.
 XX 14-AUG-2003.
 XX 21-JUN-2002; 2002WO-US019908.
 XX 22-JUN-2001; 2001US-00888164.
 XX (CELL-) CELL WORKS INC.
 XX (UYJO) UNIV JOHNS HOPKINS.
 XX Ts'o POP, Duff R, Zhou Y, Deamond S, Roby C;
 XX WPI; 2003-697456/66.
 XX New homogeneous prodrug conjugate containing hepatic ligand for delivery
 PT of pathogen-specific oligomer useful for treating liver infections or
 PT cancer.
 XX Example 2; Page 40; 107pp; English.
 XX The invention relates to a novel homogeneous conjugate comprising a
 CC hepatic ligand, bifunctional linker and biologically stable oligomer that
 CC binds to a sequence in a hepatic virus or pathogen and is released from
 CC the conjugate by hydrolysis or reduction. The conjugate of the invention
 CC may be useful during the treatment of liver diseases including chronic
 CC viral hepatitis, cirrhosis, malaria, viral or protozoan infection and
 CC cancer, such as hepatocellular carcinoma (HCC). The current sequence is
 CC that of the A-L-P conjugate-related RNA oligomer 3 of the invention.
 XX Sequence 21 BP; 5 A; 3 C; 5 G; 8 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 20; DB 10; Length 21;
 Best Local Similarity 80.0%; Pred. No. 1.2;
 Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATAAGGGTCGAUGUCCAU 20
 ::|||
 2 UUAUAAGGGUCCGAUGUCCAU 21
 Db
 RESULT 8
 ID AAT72577
 AC AAT72577 standard; DNA; 27 BP.
 XX AAT72577;
 XX 04-SEP-1997 (first entry)
 XX Hepatitis B virus RNA antisense oligonucleotide HBV94b.
 XX HBV; HBV infection; inhibition; replication; ss.
 XX Synthetic.
 XX Key
 FH Location/Qualifiers

FT misc_feature 1. 27
 FT /tag= a
 FT /note= "Internucleotide linkages are phosphorothioate"
 PN WO9639502-A1.
 XX 12-DEC-1996.
 XX 04-JUN-1996; 96WO-BP002432.
 XX 06-JUN-1995; 95US-00467397.
 XX (HOPF) HOFFMANN LA ROCHE & CO AG F.
 PA (HYBR-) HYBRIDON INC.
 XX Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
 PI Roberts NA, Roberts PC, Slade A;
 DR WPI; 1997-043124/04.
 XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
 PT used in the detection and treatment of HBV infection.
 XX Claim 1; Page 12; 81pp; English.
 XX The present sequence represents a synthetic oligonucleotide HBV94b which
 CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
 CC antisense oligonucleotide may be used to detect the presence of HBV in a
 CC sample. The antisense oligonucleotide, and oligonucleotides containing a
 CC sequence which is complementary to at least two non-contiguous regions
 CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
 CC cell or for the treatment of HBV infection
 XX Sequence 27 BP; 8 A; 4 C; 5 G; 10 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 20; DB 2; Length 27;
 Best Local Similarity 85.0%; Pred. No. 1.2;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATAAGGTCGAUGUCCAU 20
 DB 8 TTATAAGGTCGATGCCAT 27
 RESULT 9
 ID AAQ80499/c
 XX AAQ80499 standard; DNA; 28 BP.
 AC AAQ80499;
 XX 25-MAR-2003 (revised)
 DT 23-AUG-1995 (first entry)
 XX Primer to amplify hepatitis B virus core region.
 DE hepatitis B virus; X region; core region; primer; PCR; amplification;
 XX polymerase chain reaction; detection; viral infection; ss.
 XX Synthetic.
 XX WO9429483-A1.
 PN 22-DEC-1994.
 XX 03-JUN-1994; 94WO-US006360.
 XX 08-JUN-1993; 93US-00074346.
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 PA Feitelson M, Duan L, Guo J;
 PI WPI; 1995-036505/05.
 DR

XX Detection of hepatitis B virus (HBV) variants having deletions in the X
 PT region - by detection of antibodies against HBV polymerase and HB X
 PT antigen.
 XX Claim 3; Page 34; 45pp; English.
 XX This primer designated MP03 covers nucleotide bases 1903-1949 at the
 CC beginning of the hepatitis B virus (HBV) core open reading frame. It is
 CC used with MPO4 (AAQ80500) to amplify the core gene. The primers allow the
 CC detection of a specific class of HBV variants. They are useful for
 CC demonstrating the presence of productive virus infection and may prove
 CC useful in monitoring therapeutics. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 20; DB 2; Length 28;
 Best Local Similarity 85.0%; Pred. No. 1.2;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATAAGGTCGAUGUCCAU 20
 DB 20 TTATAAGGTCGATGCCAT 1
 RESULT 10
 ID AAV45201/c
 XX AAV45201 standard; DNA; 28 BP.
 AC AAV45201;
 XX 19-OCT-1998 (first entry)
 DT Primer MFO3.
 DE ss; PCR; primer; amplification; viral infection; bacterial infection;
 XX immune response; hepatitis B virus.
 XX Mus sp.
 XX WO9829121-A1.
 PN 09-JUL-1998.
 PD 02-JAN-1998; 98WO-US004116.
 XX 02-JAN-1997; 97US-0034596P.
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 PA Michaels F, Block T;
 PI WPI; 1998-387782/33.
 DR Modulating immune responses in mammals infected with infectious agent (s)
 XX - e.g. to reduce pathogenicity caused by immune responses in cases where
 PT the infectious agent has limited pathogenicity.
 XX Example 2; Page 37; 55pp; English.
 XX The primers AAV45201 and AAV45202 were used to detect the presence of a
 CC HBV genome which had been microinjected into embryos of SCID mice in an
 CC example to demonstrate modulating an immune response in a mammal infected
 CC with an infectious agent. This comprises transmucosal administration of a
 CC composition comprising an epitope which is located in close proximity to
 CC the immune response. The process may be used in treatment of mammals
 CC which are acutely or chronically infected with infectious agents, such as
 CC viruses or bacteria. It may be used to increase the immune response, or
 CC it may be used to decrease the immune response in cases where the
 CC infectious agent itself exhibits limited pathogenicity but the immune
 CC response to the infectious agent causes more significant pathogenicity.
 CC This can be the case in, e.g. hepatitis B virus (HBV) infection. The

CC process can modulate undesirable autoimmune responses exhibited by
 CC mammals infected with viral, bacterial and parasitic agents. It can
 CC prevent life-long disabilities which result from these infections
 XX

SQ Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 28;
 Best Local Similarity 85.0%; Pred. No. 1.2;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
 |||||:|:|:|:
 Db 20 TTATAAGGTCGATGCCAT 1

RESULT 11

AAA62585/c

ID AAA62585 standard; DNA; 28 BP.

XX

AC AAA62585;

XX

DT 22-NOV-2000 (first entry)

XX

DE Transgenic SCID mouse hepatitis virus transgene PCR primer MF03.

XX

KW Mouse; SCID; severe combined immunodeficiency; transgenic mouse;
 KW hepatitis virus; hepatitis B; hepatitis C; chronic liver disease;
 KW PCR primer; ss.

XX

OS Mus sp.

XX

PN US6087556-A.

XX

PD 11-JUL-2000.

XX

PF 07-JAN-1998; 98US-00003200.

XX

PR 02-MAY-1996; 96US-00641803.

XX

PA (UYJE-) UNIV JEFFERSON THOMAS.

XX

PI Siracusa L, Feitelson M;

XX

DR WPI; 2000-523731/47.

XX

PT Transgenic mouse useful in methods for evaluating interactions of
 PT chemical, drug or immunomodulating agent with hepatitis virus, lacks
 PT functional T-cells and B-cells and is capable of replicating hepatitis
 PT viruses.

XX

PS Example 3; Col 8; 7pp; English.

XX

CC The present sequence is a PCR primer used to detect a transgene in severe
 CC combined immunodeficient (SCID) mice. Transgenic immunodeficient mice
 CC were produced that are not tolerant to hepatitis viral antigens, lack
 CC functional T-cells and B-cells and contain integrated hepatitis virus DNA
 CC in the somatic and germ cells. The hepatitis virus gene is expressed and
 CC the hepatitis virus is replicated in the transgenic mouse. The mouse may
 CC be used as an animal model for evaluating interactions of a chemical,
 CC drug or immunomodulating agent with a hepatitis virus. It is also useful
 CC for the assessment of anti-viral and immunomodulatory intervention
 CC therapies, including the screening of drug candidates. It can be used to
 CC analyse the virus-host relationship, to evaluate the relationship between
 CC the virus and chemicals metabolised and/or detoxified by the liver, and
 CC to identify cellular biochemical pathways contributing to the development
 CC and progression of chronic liver disease. The transgenic mouse is thus
 CC useful for elucidating the effects of hepatitis virus on hepatic
 CC metabolism

XX

SQ Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 28;
 Best Local Similarity 85.0%; Pred. No. 1.2;

Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATAAGGTCGAUGUCCAU 20
 |||||:|:|:|:
 Db 20 TTATAAGGTCGATGCCAT 1

RESULT 12

ADN36070/c

ID ADN36070 standard; DNA; 28 BP.

XX

AC ADN36070;

XX

DT 01-JUL-2004 (first entry)

XX

DE Probe #151 to determine effect of long term lamivudine treatment of HBV.

XX

KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.

XX

OS Hepatitis B virus.

XX

PN WO2004031729-A2.

XX

PD 15-APR-2004.

XX

PF 01-OCT-2003; 2003WO-US031121.

XX

PR 01-OCT-2002; 2002US-0415301P.

XX

PA (GEOU) UNIV GEORGETOWN.

XX

PI Korba BB, Ciancio A, Gerin JL;

XX

DR WPI; 2004-348004/32.

XX

PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
 PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
 PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.

XX

PS Claim 31; SEQ ID NO 151; 107pp; English.

XX

CC The invention relates to a method of predicting the long term response of
 CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
 CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
 CC leucine at amino acid position (aa) 91 in the DNA polymerase region
 CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
 CC 604) in the DNA polymerase region of HBV. The method comprises
 CC determining whether the HBV carried by the host bears one or more of the
 CC following mutations: (i) Q213S (Glutamine to serine at aa213) (originally
 CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
 CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
 CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
 CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
 CC and C1962A representing specific double point mutations in the DNA
 CC precore/core promoter or ORF region. The method and kit is useful in
 CC predicting the long-term response of a host of HBV to 3TC therapy (also
 CC known as lamivudine). This sequence represents an oligonucleotide
 CC sequence used in the method of the invention to detect a mutation in the
 CC above mentioned sequences.

SQ Sequence 28 BP; 11 A; 6 C; 5 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;

Best Local Similarity 85.0%; Pred. No. 1.2;

Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20

|||||:|:|:|:

Db 23 TTATAAGGTCGATGCCAT 4

RESULT 13

ADN36068/c

```

ID ADN36068 standard; DNA; 28 BP.
XX AC ADN36068;
XX DT 01-JUL-2004 (first entry)
XX DE Probe #149 to determine effect of long term lamivudine treatment of HBV.
XX KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX OS Hepatitis B virus.
XX PN WO2004031729-A2.
XX PD 15-APR-2004.
XX PF 01-OCT-2003; 2003WO-US031121.
XX PR 01-OCT-2002; 2002US-0415301P.
XX PA (GEOU ) UNIV GEORGETOWN.
XX PI Korba BE, Ciancio A, Gerin JL;
XX DR WPI; 2004-348004/32.
XX PF Predicting the long-term response of a host of hepatitis B virus (HBV) to
XX PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
XX PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX PS Claim 31; SEQ ID NO 149; 107pp; English.
XX CC The invention relates to a method of predicting the long term response of
XX CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
XX CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
XX CC leucine at amino acid position (aa) 91 in the DNA polymerase region
XX CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
XX CC 604) in the DNA polymerase region of HBV. The method comprises
XX CC determining whether the HBV carried by the host bears one or more of the
XX CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
XX CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
XX CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
XX CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
XX CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
XX CC and C1962A representing specific double point mutations in the DNA
XX CC precore/core promoter or ORF region. The method and kit is useful in
XX CC predicting the long-term response of a host of HBV to 3TC therapy (also
XX CC known as lamivudine). This sequence represents an oligonucleotide
XX CC sequence used in the method of the invention to detect a mutation in the
XX CC above mentioned sequences.
XX SQ Sequence 28 BP; 11 A; 6 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 85.0%; Pred. No. 1.2;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAGGGTTCGAUGUCCAU 20
Db |||||:|||||:|||||:|||||:
25 TTATAGGGTTCGATGCCAT 6

RESULT 14
ADN36071/c
ID ADN36071 standard; DNA; 28 BP.
XX AC ADN36071;
XX DT 01-JUL-2004 (first entry)
XX DE Probe #152 to determine effect of long term lamivudine treatment of HBV.
XX KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.

```

```

XX OS Hepatitis B virus.
XX PN WO2004031729-A2.
XX PD 15-APR-2004.
XX PF 01-OCT-2003; 2003WO-US031121.
XX PR 01-OCT-2002; 2002US-0415301P.
XX PA (GEOU ) UNIV GEORGETOWN.
XX PI Korba BE, Ciancio A, Gerin JL;
XX DR WPI; 2004-348004/32.
XX PF Predicting the long-term response of a host of hepatitis B virus (HBV) to
XX PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
XX PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX PS Claim 31; SEQ ID NO 152; 107pp; English.
XX CC The invention relates to a method of predicting the long term response of
XX CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
XX CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
XX CC leucine at amino acid position (aa) 91 in the DNA polymerase region
XX CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
XX CC 604) in the DNA polymerase region of HBV. The method comprises
XX CC determining whether the HBV carried by the host bears one or more of the
XX CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
XX CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
XX CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
XX CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
XX CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
XX CC and C1962A representing specific double point mutations in the DNA
XX CC precore/core promoter or ORF region. The method and kit is useful in
XX CC predicting the long-term response of a host of HBV to 3TC therapy (also
XX CC known as lamivudine). This sequence represents an oligonucleotide
XX CC sequence used in the method of the invention to detect a mutation in the
XX CC above mentioned sequences.
XX SQ Sequence 28 BP; 11 A; 6 C; 4 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 85.0%; Pred. No. 1.2;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATAGGGTTCGAUGUCCAU 20
Db |||||:|||||:|||||:|||||:
22 TTATAGGGTTCGATGCCAT 3

RESULT 15
ADN36073/c
ID ADN36073 standard; DNA; 28 BP.
XX AC ADN36073;
XX DT 01-JUL-2004 (first entry)
XX DE Probe #154 to determine effect of long term lamivudine treatment of HBV.
XX KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
XX OS Hepatitis B virus.
XX PN WO2004031729-A2.
XX PD 15-APR-2004.
XX PF 01-OCT-2003; 2003WO-US031121.
XX

```

```

PR 01-OCT-2002; 2002US-0415301P.
XX (GEOU ) UNIV GEORGETOWN.
PA
XX Korba BE, Cíancio A, Gerin JL;
XX WPI; 2004-348004/32.
XX
XX Predicting the long-term response of a host of hepatitis B virus (HBV) to
PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
XX
XX Claim 31; SEQ ID NO 154; 107pp; English.
XX
XX The invention relates to a method of predicting the long term response of
XX a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
XX the HBV carried by the host (i) bears a nucleic acid that encodes for a
XX leucine at amino acid position (aa) 91 in the DNA polymerase region
XX (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
XX 604) in the DNA polymerase region of HBV. The method comprises
XX determining whether the HBV carried by the host bears one or more of the
XX following mutations: (i) Q213S (glutamine to serine at aa213) (originally
XX codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
XX T1960G, or T1961A/G specific point mutation in the DNA precore/core
XX promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
XX changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
XX and C1962A representing specific double point mutations in the DNA
XX precore/core promoter or ORF region. The method and kit is useful in
XX predicting the long-term response of a host of HBV to 3TC therapy (also
XX known as lamivudine). This sequence represents an oligonucleotide
XX sequence used in the method of the invention to detect a mutation in the
XX above mentioned sequences.
XX
SQ Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match      100.0%; Score 20; DB 12; Length 28;
Best Local Similarity 85.0%; Pred. NO. 1.2;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTATAGGGTCGAUGUCCAU 20
DB      20 TTATAGGGTCGATGCCAT 1

```

Search completed: March 17, 2005, 06:48:44
Job time : 172.333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612A-64
Perfect score: 20
Sequence: 1 ttataagggtcgauuccau 20
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.4	92.0	706	4	BI888237
C 2	17.4	87.0	738	8	AZ195485 SP 1030.A
C 3	16.8	84.0	209	4	BM024704 fu71e06.x
C 4	16.8	84.0	211	6	CD014361 hac31a04.
C 5	16.8	84.0	331	9	CR405240 Arabidops
C 6	16.8	84.0	420	7	H98675 yx17d02.sl
C 7	16.8	84.0	426	4	BM571757 fx05f04.x
C 8	16.8	84.0	433	4	BI840986 fg41f02.x
C 9	16.8	84.0	438	2	AW147602 gal3f06.y
C 10	16.8	84.0	475	7	H98688 yx17h02.sl
C 11	16.8	84.0	477	4	BM571990 fx05f04.y
C 12	16.8	84.0	495	5	BM093445 BM093445
C 13	16.8	84.0	508	4	BM072332 fu08d10.x
C 14	16.8	84.0	528	5	BM234048 BM234048
C 15	16.8	84.0	551	5	BM392185 BM392185
C 16	16.8	84.0	553	4	BM185735 fu71e09.x
C 17	16.8	84.0	562	4	BM185703 fu71a05.x
C 18	16.8	84.0	563	4	BM005095 fu64c11.x
C 19	16.8	84.0	580	4	BI983042 fu42f10.x
C 20	16.8	84.0	584	4	BM005060 fu63g11.x
C 21	16.8	84.0	586	4	BM024740 fu72b11.x
C 22	16.8	84.0	586	9	CR333657 Medicago
C 23	16.8	84.0	600	5	BM338052 BM338052
C 24	16.8	84.0	621	6	CD014379 hac31c04.

25	16.8	84.0	695	6	CD237681
C 26	16.8	84.0	710	5	BM018497
C 27	16.8	84.0	711	9	BM222138
C 28	16.8	84.0	714	8	AQ648647 RFC193-Bc
C 29	16.8	84.0	729	4	BJ705222 BJ705222
C 30	16.8	84.0	733	1	AI635930 t282c11.x
C 31	16.8	84.0	734	9	BI138600 Danio rer
C 32	16.8	84.0	738	9	BM245774 Danio rer
C 33	16.8	84.0	774	5	BM025653 BM025653
C 34	16.8	84.0	775	9	BI170776 Danio rer
C 35	16.8	84.0	780	9	BI144130 Danio rer
C 36	16.8	84.0	796	5	BM023432 BM023432
C 37	16.8	84.0	821	7	CK129749 ACENECOURT
C 38	16.8	84.0	860	9	CG002233 ZUABR83TH
C 39	16.4	82.0	413	5	BP569691 BP569691
C 40	16.4	82.0	492	9	CG596834 OST259526
C 41	16.4	82.0	573	9	CG594911 OST253619
C 42	16.4	82.0	584	7	CK529842 rswfa0_01
C 43	16.4	82.0	597	7	CK529191 rswfa0_00
C 44	16.4	82.0	623	7	CK528881 rswfa0_00
C 45	16.4	82.0	656	5	BM362244 BM362244

ALIGNMENTS

RESULT 1
BI888237/c 706 bp mRNA linear EST 12-OCT-2001
LOCUS ZF637-1-002492 Zebrafish shield stage whole embryo cDNA library
DEFINITION MPMPGp637 Danio rerio cDNA clone MPMPGp637_18P2;MPMPGp637P0218 5', mRNA sequence.
ACCESSION BI888237
VERSION BI888237.1 GI:16095508
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1. (bases 1 to 706)
AUTHORS Clark, M., Aanstad, P., Hennig, S., Johnson, S.L. and Lehrach, H.
TITLE EST sequencing of a zebrafish shield stage cDNA library normalised by oligonucleotide fingerprinting
JOURNAL Unpublished (2001)
COMMENT Contact: Hennig S
Laboraty 123, dept.Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
5' EST sequencing of clones from a zebrafish shield stage library, normalised from 55,000 starting clones by oligonucleotide fingerprinting
High quality sequence stop: 706.
Location/Qualifiers
1. .706
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MPMPGp637_18P2;MPMPGp637P0218"
/issue_type="whole embryo"
/lab_host="E.coli XL1 blue MRF"
/clone_lib="zebrafish shield stage whole embryo cDNA library MPMPGp637"
/notes="Vector: pSport1; Site 1: NotI; Site 2: SalI; oligo-dT-NotI primed, SalI adaptors, directionally cloned, library normalised by oligonucleotide fingerprinting"

Query Match 92.0%; Score 18.4; DB 4; Length 706;

```

Best Local Similarity 80.0%; Pred. No. 55;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
Db 499 TTATAAGGTCGATGTCAT 480

RESULT 2
AZ195485
LOCUS
DEFINITION
SP 1030 Al H09 SP6E Strongylocentrotus purpuratus, purple sea
urchin, sperm genomic BAC library Strongylocentrotus purpuratus
genomic clone Plate=1030 Col=17 Row=O, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
REFERENCE
1 (bases 1 to 738)
Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T.,
Wray,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H.
and Hood,L.
A sea urchin genome project: Sequence scan, virtual map, and
additional resources
JOURNAL
MEDLINE
PUBMED
20402566
10920195
COMMENT
Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 1030 row: O column: 17
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 738.
Location/Qualifiers
1..738
/organism="Strongylocentrotus purpuratus"
/mol_type="genomic DNA"
/db_xref="taxon:7668"
/clone_lib="Plate=1030 Col=17 Row=O"
/urchin, sperm genomic BAC library"
/notes="Organ: sperm; Vector: BACe3.6; BAC Clones in E-Coli
DH10B"

ORIGIN
Query Match 87.0%; Score 17.4; DB 8; Length 738;
Best Local Similarity 78.9%; Pred. No. 1.9e+02;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 TATAAGGTCGAUGUCCAU 20
Db 9 TATAAGGGCGCATGTCAT 27

RESULT 3
BM024704
LOCUS
DEFINITION
fu7le06.x1 zebrafish adult brain panio rerio cDNA clone
IMAGE:5334995 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
BM024704.1 GI:116539060
Danio rerio (zebrafish)

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 209)
Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterson,R. and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
RessourcenzentrumPrimardatenbank, Berlin, Germany (web address:
www.rzpd.de)
Seq primer: -40UP
High quality sequence stop: 200.
Location/Qualifiers
1..209
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5334995"
/sex="mixed male and female"
/tissue_type="brain"
/dev_stage="adult"
/lab_host="E. coli DH10B"
/clone_lib="zebrafish adult brain"
/notes="Vector: pZIPLOX; Site_1: NotI; Site_2: SalI;
Original library was constructed in lambdaZ1PLOX. Mass
excision of the cDNA library was performed to yield
pZIPLOX plasmids. Insert check was done in original
library."

ORIGIN
Query Match 84.0%; Score 16.8; DB 4; Length 209;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
Db 176 TTATAAGGTCGATGTCAT 195

RESULT 4
CD014361
LOCUS
DEFINITION
hac3la04.x1 MPZFRIken1 Danio rerio cDNA clone IMAGE:6923625 3',
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
CD014361.1 GI:30330820
Danio rerio (zebrafish)

ORIGIN
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 211)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,

```

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 WashU Zebrafish EST Project 1998
 Unpublished (1998)
 JOURNAL
 COMMENT
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrafish@watson.wustl.edu
 Oligo-dt primed cDNA was produced from intestine (two independent samples) and liver using
 5'-GAGAGAGAGAGATCCAAAXXXXXTTTTTTTTTTTTTTTNN-3' with the following tags substituting for XXXXX: CAGAG (tag C04): liver; CGGTAT (tag C12): intestine; and CGTATG (tag D01): intestine. cDNA was pooled and size- selected for a 1.6 kb average insert. Library was constructed using the Cap- Trapper method as described in Genomics 2001: 77(1-2)79-90. Library donated by M. Pack, M.D. (University of Pennsylvania School of Medicine).
 Seq primer: T3 from Gibco
 High quality sequence stop: 211.
 Location/Qualifiers
 1. 211
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:6923625"
 /tissue_type="intestine (2 samples) and liver (pooled)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="MPZFRikenl"
 /notes="Vector: pBluescriptR; Site 1: XhoI; Site 2: BamHI; Oligo-dt primed cDNA was produced from intestine (two independent samples) and liver using
 5'-GAGAGAGAGAGATCCAAAXXXXXTTTTTTTTTTTTTTTNN-3' with the following tags substituting for XXXXX: CAGAG (tag C04): liver; CGGTAT (tag C12): intestine; and CGTATG (tag D01): intestine. cDNA was pooled and size- selected for a 1.6 kb average insert. Library was constructed using the Cap- Trapper method as described in Genomics 2001: 77(1-2)79-90. Library donated by M. Pack, M.D. (University of Pennsylvania School of Medicine)."
 ORIGIN
 Query Match 84.0%; Score 16.8; DB 6; Length 211;
 Best Local Similarity 75.0%; Pred. No. 3.5e+02;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TTATAGGGTCGAUGUCCAU 20
 |||||:|||||:|:|:
 Db 32 TTATTAGGTCGATGTGCAT 51
 CR405240 331 bp DNA linear GSS 02-MAY-2004
 LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-876H09-026468,
 DEFINITION genomic survey sequence.
 CR405240
 VERSION CR405240.1 GI:46945968
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1
 Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P. and Weissshaar, B.
 GABI-kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
 Bioinformatics 19 (11), 1441-1442 (2003)
 JOURNAL
 MEDLINE 22755829

PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source

12874060
 2 Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weissshaar, B.
 An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
 Plant Mol. Biol. 53 (1-2), 247-259 (2003)
 23117147
 PUBMED 14756321
 3 Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and Weissshaar, B.
 High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
 BioTechniques 35 (6), 1164-1168 (2003)
 14682050
 4 (bases 1 to 331)
 Rosso, M.G., Strizhov, N., Li, Y. and Weissshaar, B.
 Direct Submission
 Submitted (01-MAY-2004) Weissshaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
 This sequence has been recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by BAC clone F9D12. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
 Location/Qualifiers
 1. 331
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="GK-876H09-026468"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /ecotype="Col-0"
 /note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC161 (Genbank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion.
 T-DNA derived sequences were removed."
 Query Match 84.0%; Score 16.8; DB 9; Length 331;
 Best Local Similarity 75.0%; Pred. No. 3.7e+02;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TTATAGGGTCGAUGUCCAU 20
 |||||:|||||:|:|:
 Db 35 TTATAGGGAGGATGTCCAT 16
 RESULT 6
 H98675
 LOCUS Yx17d02.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
 DEFINITION IMAGE:261987 3', mRNA sequence.
 H98675
 VERSION H98675.1 GI:1123343
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 420)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Huitman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and

TITLE
JOURNAL
COMMENT

Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 237
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: ml3 -40 forward
High quality sequence stop: 237.

FEATURES

source

1..420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3871629"
/db_xref="taxon:9606"
/clone="IMAGE:261987"
/sex="Male"
/tissue_type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares melanocyte 2MbHM"
/notes="Vector: pT7T3D (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCGAATGGAGCGCCGAGTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bonaldo. RNA from normal foreskin melanocytes
(FS374) was kindly provided by Dr. Anthony P. Albino."

ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 420;
Best Local Similarity 75.0%; Pred. No. 3.8e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20

Db 94 TTCTCAGGTCGATGCCAT 113

RESULT 7

BM571757

LOCUS

DEFINITION BM571757 426 bp mRNA linear EST 21-FEB-2002
IMAGE:5619198 3', mRNA sequence.

ACCESSION BM571757.1 GI:18853740

VERSION

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 426)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,

Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,

Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,

Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,

Waterston, R. and Wilson, R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

TITLE

JOURNAL

COMMENT

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

FEATURES

source

1..426
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5619198"
/sex="female"
/dev_stage="4-5 month"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Gong zebrafish ovary"
/note="Organ: ovary (pooled); Vector: pBluescript SK-;
Site 1: XhoI; Site 2: EcoRI; Poly A+ RNA was isolated from
the ovaries of 2 female adult zebrafish (4-5 month old).
cDNAs were made using oligo-dT primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
mass-excised to pBluescript SK- following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Zhiyuan Gong for further information on
this library (National University of Singapore,
Department of Biological Sciences, Lower Kent Ridge Road,
Singapore 119260)."

ORIGIN

Query Match 84.0%; Score 16.8; DB 4; Length 426;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20

Db 205 TTATTAGGTCGATGCCAT 224

RESULT 8

BI840986

LOCUS

DEFINITION

BI840986

VERSION

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 433)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,

Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,

Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,

Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,

Waterston, R. and Wilson, R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu

cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:

Matthew Clark. DNA Sequencing by: Washington University Genome

Sequencing Center Clone distribution: Genome Systems, St. Louis,

Missouri (web address: www.genomesystems.com) (email contact:

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu

The library was constructed by Dr. Z. Gong. DNA Sequencing by:
Washington University Genome Sequencing Center St. Louis. Please
contact Zhiyuan Gong for further information on this library
(National University of Singapore, Department of Biological
Sciences, Lower Kent Ridge Road, Singapore 119260).
Seq primer: T7 from Gibco.

Location/Qualifiers

info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and ResourcenZentrumPrimarDatenbank, Berlin, Germany (web address: www.zpdp.de)

Seq primer: -40UP
High quality sequence stop: 396.

FEATURES

source

```

1. .433
  Location/Qualifiers
  /organism="Danio rerio"
  /mol_type="mRNA"
  /db_xref="taxon:7955"
  /clone="IMAGE:483393"
  /sex="mixed male and female"
  /tissue_type="brain"
  /dev_stage="adult"
  /lab_host="E. coli DH10B"
  /clone_lib="zebrafish adult brain"
  /note="Vector: pZIPLOX; Site 1: NotI; Site 2: SalI;
  Original library was constructed in lambdaZAPLOX. Mass
  excision of the cDNA library was performed to yield
  pZIPLOX plasmids. Insert check was done in original
  library."

```

ORIGIN

Query Match 84.0%; Score 16.8; DB 4; Length 433;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAGGGTCGAUGUCCAU 20

Db 182 TTATAGGGTCGATGTCAT 201

RESULT 9

AW147602/c

LOCUS

DEFINITION AW147602 438 bp mRNA linear EST 19-FEB-2003
Gai3f06.y1 normalized Xenopus laevis gastrula cDNA
clone XENOPUS SOURCE ID:xlnga001c12 5' similar to SW:AOP2_HUMAN
P30041 ANTIOXIDANT PROTEIN 2 ; mRNA sequence.

ACCESSION AW147602

VERSION AW147602.1 GI:6195498

KEYWORDS

SOURCE EST.

ORGANISM

Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE 1 (bases 1 to 438)

Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L.,
Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
Person,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
Waterston,R. and Wilson,R.

WashU Xenopus EST project, 1999

Unpublished (1999)

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by Bruce Blumberg

Library normalized by Jihwan Song

DNA Sequencing by: Washington University Genome Sequencing Center

Seq primer: -40RP from Gibco

High quality sequence stop: 385.

FEATURES

source

```

1. .438
  Location/Qualifiers
  /organism="Xenopus laevis"
  /mol_type="mRNA"
  /db_xref="taxon:8355"
  /clone="XENOPUS_SOURCE ID:xlnga001c12"
  /tissue_type="gastrula (stages 10.5, 11.5 mixed)"

```

/lab_host="Top-10 F"
/clone_lib="normalized Xenopus laevis gastrula"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from 2ug of poly A+ RNA (equal
parts from stage 10.5 and stage 11.5 gastrulae).
EcoRI-XhoI cut cDNA was then ligated into UniZap-XR
(Stratagene) with EcoRI at the 5' end and XhoI at the 3'
end. SS-library phagmids were prepared by mass excision
from the original library and normalized by hybridization
to biotinylated driver (prepared from the same library by
PCR) to Cot-omega of 11. After removal of hybrids and
excess driver by streptavidin sepharose chromatography,
the ss-phagmids were made double stranded and
electroporated into Top-10 F'. Original library
construction by Bruce Blumberg (Cho et al. 1991 Cell 67,
1111-1120). Normalized by Jihwan Song (Song, Cho and
Blumberg, unpublished). Note: This is a Xenopus Gene
Collection (XGC) library."

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 438;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAGGGTCGAUGUCCAU 20

Db 299 TTATAGAGTTGATGCCAT 280

RESULT 10

H98688

LOCUS

DEFINITION yx17h02.s1 Soares melanocyte 2NBHM Homo sapiens cDNA clone
IMAGE:262035 3', mRNA sequence.

ACCESSION H98688

VERSION H98688.1 GI:1123356

KEYWORDS

SOURCE EST.

ORGANISM

Homo sapiens (human)
Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 475)

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisan,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.

The WashU-Merck EST Project

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 324

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: ml3 -40 forward

High quality sequence stop: 324.

FEATURES

source

```

1. .475
  Location/Qualifiers
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="GBB:3871677"
  /db_xref="taxon:9606"
  /clone="IMAGE:262035"
  /sex="Male"
  /tissue_type="melanocyte"
  /lab_host="DH10B (ampicillin resistant)"
  /clone_lib="Soares melanocyte 2NBHM"
  /note="Vector: pT73D (Pharmacia) with a modified

```

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGAGTGTGTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pMT3 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."

ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 475;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
|||||:|:|:|:
Db 93 TTCTCAGGTCGATGTCAT 112

RESULT 11

BM571990/c
LOCUS
DEFINITION 477 bp mRNA linear EST 21-FEB-2002
IMAGE:5619198 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

Danio rerio (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE

AUTHORS
Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M.,
Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Stepec, M., Theising, S., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
WashU Zebrafish EST Project 1998

TITLE

Unpublished (1998)

JOURNAL

COMMENT

Other_ESTs: fx05f04.xl
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu

The library was constructed by Dr. Z. Gong. DNA Sequencing by:
Washington University Genome Sequencing Center St. Louis. Please
contact Zhiyuan Gong for further information on this library
(National University of Singapore, Department of Biological
Sciences, Lower Kent Ridge Road, Singapore 119260).
Seq primer: T3 ET from Amerham
High quality sequence stop: 394.

FEATURES

source

1..477
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5619198"
/sex="female"
/dev_stage="4-5 month"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Gong zebrafish ovary"

/note="Organ: ovary (pooled); Vector: pBluescript SK-;
Site 1: XhoI; Site 2: EcoRI; Poly A+ RNA was isolated from
the ovaries of 2 female adult zebrafish (4-5 month old).
cDNAs were made using oligo-dT primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
mass-excised to pBluescript SK- following the Washington
University protocol

ORIGIN

Query Match 84.0%; Score 16.8; DB 4; Length 477;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
|||||:|:|:|:
Db 291 TTATTAGGTCGATGTCAT 272

RESULT 12

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

source

```

REFERENCE
AUTHORS      1 (bases 1 to 508)
              Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
              Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
              Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
              Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
              Ritter,B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
              Waterston,R. and Wilson,R.
TITLE        WashU Zebrafish EST Project 1998
JOURNAL      Unpublished (1998)
COMMENT      Contact: Stephen L. Johnson
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: zbrfish@watson.wustl.edu
              cDNA Library Preparation: John Ngai. cDNA Library Arrayed by:
              Matthew Clark. DNA Sequencing by: Washington University Genome
              Sequencing Center Clone distribution: Genome Systems, St. Louis,
              Missouri (web address: www.genomesystems.com) (email contact:
              info@genomesystems.com) and Research Genetics, Huntsville, Alabama
              (web address: www.resgen.com) (email contact: info@resgen.com) and
              RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
              www.rzpd.de)
              Seq primer: -40UP
              High quality sequence stop: 444.
FEATURES
source       1..508
              Location/Qualifiers
              /organism="Danio rerio"
              /mol_type="mRNA"
              /db_xref="taxon:7955"
              /clone="IMAGE:5386027"
              /sex="mixed male and female"
              /tissue_type="brain"
              /dev_stage="adult"
              /lab_host="E. coli DH10B"
              /clone_lib="zebrafish adult brain"
              /notes="vector: pZiPlox; Site 1: NotI; Site 2: SalI;
              Original library was constructed in lambdaZiPlox. Mass
              excision of the cDNA library was performed to yield
              pZiPlox plasmids. Insert check was done in original
              library."
ORIGIN
Query Match      84.0%; Score 16.8; DB 4; Length 508;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAGGTCGAUGUCCAU 20
    ||||| ||||| ||||| |||||
DB 180 TTATTAGGTCGATGTCAT 199

RESULT 14
BW234048
LOCUS          528 bp mRNA linear EST 07-NOV-2002
DEFINITION    BW234048 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
               intestinalis cDNA clone citb046h04 5', mRNA sequence.
ACCESSION     BW234048.1 GI:24755889
VERSION       BW234048
KEYWORDS      EST.
SOURCE        Ciona intestinalis
ORGANISM      Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
               Phlebobranchia; Cionidae; Ciona.
REFERENCE     1 (bases 1 to 528)
AUTHORS       Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE         Expressed genes in Ciona intestinalis (2002c)
JOURNAL       Unpublished (2002)
COMMENT       Contact: Nori Satoh
               Department of Zoology
               Kyoto University
               Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
               Tel: 81-75-753-4081

Query Match      84.0%; Score 16.8; DB 4; Length 508;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAGGTCGAUGUCCAU 20
    ||||| ||||| ||||| |||||
DB 180 TTATTAGGTCGATGTCAT 199

REFERENCE
AUTHORS      1 (bases 1 to 528)
              Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE        Expressed genes in Ciona intestinalis (2002c)
JOURNAL      Unpublished (2002)
COMMENT      Contact: Nori Satoh
              Department of Zoology
              Kyoto University
              Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
              Tel: 81-75-753-4081

```

```

Fax: 81-75-705-1113
Email: sath@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
source       1..528
              /organism="Ciona intestinalis"
              /mol_type="mRNA"
              /db_xref="taxon:7719"
              /clone="citb046h04"
              /tissue_type="whole animal"
              /dev_stage="tailbud embryo"
              /clone_lib="Nori Satoh unpublished cDNA library, tailbud
              embryo"
ORIGIN
Query Match      84.0%; Score 16.8; DB 5; Length 528;
Best Local Similarity 75.0%; Pred. No. 4e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
    ||||| ||||| ||||| |||||
DB 160 TTATAAGTCGTGATGCCAT 179

RESULT 15
BW392185/c
LOCUS          551 bp mRNA linear EST 28-MAY-2004
DEFINITION    BW392185 Yutaka Satou unpublished cDNA library, embryo whole animal
               Ciona intestinalis cDNA clone ciem807o10 3', mRNA sequence.
ACCESSION     BW392185
VERSION       BW392185.1 GI:47808013
KEYWORDS      EST.
SOURCE        Ciona intestinalis
ORGANISM      Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
               Phlebobranchia; Cionidae; Ciona.
REFERENCE     1 (bases 1 to 551)
AUTHORS       Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE         Expressed genes in Ciona intestinalis (2004)
JOURNAL       Unpublished (2004)
COMMENT       Contact: Yutaka Satou
               Department of Zoology
               Kyoto University
               Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
               Tel: 81-75-753-4095
               Fax: 81-75-705-1113
               Email: yutaka@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
source       1..551
              /organism="Ciona intestinalis"
              /mol_type="mRNA"
              /db_xref="taxon:7719"
              /clone="ciem807o10"
              /tissue_type="whole animal"
              /dev_stage="embryo"
              /clone_lib="Yutaka Satou unpublished cDNA library, embryo
              whole animal"
ORIGIN
Query Match      84.0%; Score 16.8; DB 5; Length 551;
Best Local Similarity 75.0%; Pred. No. 4e+02;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTATAAGGTCGAUGUCCAU 20
    ||||| ||||| ||||| |||||
DB 319 TTATAAGTCGTGATGCCAT 300

Search completed: March 17, 2005, 11:07:52
Job time : 1389.27 secs

```

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:07:57 ; Search time 683.733 Seconds
(without alignments)
1417.372 Million cell updates/sec

Title: US-08-901-612A-65
Perfect score: 20
Sequence: 1 aaattctttataagggucca 20
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_uni:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6 AR027821	AR027821 Sequence
2	20	100.0	27	6 AR027819	AR027819 Sequence
C 3	20	100.0	28	6 AR103926	AR103926 Sequence
C 4	20	100.0	35	6 BD236992	BD236992 DNA vacci
C 5	20	100.0	81	6 I92348	I92348 Sequence 9
C 6	20	100.0	253	14 AY329529	AY329529 Hepatitis
C 7	20	100.0	253	14 AY329561	AY329561 Hepatitis
C 8	20	100.0	253	14 AY329562	AY329562 Hepatitis
C 9	20	100.0	253	14 AY329568	AY329568 Hepatitis
C 10	20	100.0	253	14 AY329573	AY329573 Hepatitis
C 11	20	100.0	253	14 AY329575	AY329575 Hepatitis
C 12	20	100.0	253	14 AY329581	AY329581 Hepatitis
C 13	20	100.0	294	14 AF390000	AF390000 Hepatitis
C 14	20	100.0	333	14 HPBHBED	L12359 Hepatitis B
C 15	20	100.0	398	14 AB167603	AB167603 Hepatitis
C 16	20	100.0	398	14 AB167637	AB167637 Hepatitis
C 17	20	100.0	406	14 AB163815	AB163815 Hepatitis
C 18	20	100.0	406	14 AB163817	AB163817 Hepatitis
C 19	20	100.0	439	14 AY254503	AY254503 Hepatitis

C 20	100.0	456	14	AY509204	Hepatitis
C 21	20	100.0	488	14 AY274419	Hepatitis
C 22	20	100.0	488	14 AY274420	Hepatitis
C 23	20	100.0	488	14 AY274422	Hepatitis
C 24	20	100.0	488	14 AY274424	Hepatitis
C 25	20	100.0	488	14 AY274427	Hepatitis
C 26	20	100.0	488	14 AY274428	Hepatitis
C 27	20	100.0	488	14 AY274429	Hepatitis
C 28	20	100.0	488	14 AY274430	Hepatitis
C 29	20	100.0	488	14 AY274431	Hepatitis
C 30	20	100.0	488	14 AY274432	Hepatitis
C 31	20	100.0	488	14 AY274433	Hepatitis
C 32	20	100.0	488	14 AY274434	Hepatitis
C 33	20	100.0	488	14 AY274436	Hepatitis
C 34	20	100.0	548	14 AY382500	Hepatitis
C 35	20	100.0	548	14 AY382501	Hepatitis
C 36	20	100.0	548	14 AY382502	Hepatitis
C 37	20	100.0	548	14 AY382521	Hepatitis
C 38	20	100.0	548	14 AY382522	Hepatitis
C 39	20	100.0	548	14 AY382523	Hepatitis
C 40	20	100.0	548	14 AY382524	Hepatitis
C 41	20	100.0	548	14 AY382525	Hepatitis
C 42	20	100.0	548	14 AY382526	Hepatitis
C 43	20	100.0	548	14 AY382527	Hepatitis
C 44	20	100.0	552	6 BD236991	DNA vacci
C 45	20	100.0	552	14 AB023678	Hepatitis

ALIGNMENTS

RESULT 1
LOCUS AR027821 20 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 19 from patent US 5856459.
ACCESSION AR027821
VERSION AR027821.1 GI:5938641
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and Mills,J.S.
TITLE Oligonucleotides specific for hepatitis B virus
JOURNAL Patent: US 5856459-A 19 05-JAN-1999;
FEATURES Location/Qualifiers
source 1..20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAGGGUCCA 20
|||||
Db 1 AAATCTTTTATAGGGTCCA 20

RESULT 2
LOCUS AR027819 27 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 17 from patent US 5856459.
ACCESSION AR027819
VERSION AR027819.1 GI:5938639
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Frank,B.L., Roberts,P.C., Goodchild,J., Craig,J.Charles. and

```

Mills,J.S.
TITLE      Oligonucleotides specific for hepatitis B virus
JOURNAL    Patent: US 5856459-A 17 05-JAN-1999;
FEATURES   Location/Qualifiers
            source
            1..27
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 27;
Best Local Similarity 95.0%; Pred. No. 1.3e+02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAAGGUCGA 20
    |||||
Db 1 AAATCTTTTATAAGGTCGA 20

RESULT 3
AR103926/c
LOCUS       AR103926             28 bp      DNA      linear      PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6087556.
ACCESSION  AR103926
VERSION     AR103926.1 GI:12815514
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 28)
AUTHORS   Feitelson,M. and Siracusa,L.
TITLE     Transgenic animals capable of replicating hepatitis viruses and
           mimicking chronic hepatitis infection in humans
JOURNAL   Patent: US 6087556-A 11-JUL-2000;
FEATURES   Location/Qualifiers
            source
            1..28
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 28;
Best Local Similarity 95.0%; Pred. No. 1.3e+02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAAGGUCGA 20
    |||||
Db 27 AAATCTTTTATAAGGTCGA 8

RESULT 4
BD236992/c
LOCUS       BD236992             35 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION DNA vaccination to cholesterol ester transfer protein in the
           treatment of atherosclerosis.
ACCESSION  BD236992
VERSION     BD236992.1 GI:33046762
KEYWORDS   JP 2002516656-A/17.
SOURCE     unidentified
ORGANISM   unidentified.
REFERENCE  1 (bases 1 to 35)
AUTHORS   Needleman,P. and Glenn,K.
TITLE     DNA vaccination to cholesterol ester transfer protein in the
           treatment of atherosclerosis
JOURNAL   Patent: JP 2002516656-A 17 11-JUN-2002;
COMMENT    MONSANTO CO
           OS   Unidentified
           PN   JP 2002516656-A/17
           PD   11-JUN-2002
           PE   17-SEP-1998 JP 2000512947
           PR   19-SEP-1997 US 08/934367
           PI   PHILIP NEEDLEMAN,KEVIN GLENN
           PC   C12N15/09,A61K48/00,C12N15/00
           CC   Strandedness: Single;

Mills,J.S.
TITLE      Oligonucleotides specific for hepatitis B virus
JOURNAL    Patent: US 5856459-A 17 05-JAN-1999;
FEATURES   Location/Qualifiers
            source
            1..35
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 35;
Best Local Similarity 95.0%; Pred. No. 1.3e+02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAAGGUCGA 20
    |||||
Db 33 AAATCTTTTATAAGGTCGA 14

RESULT 5
I92348/c
LOCUS       I92348             81 bp      DNA      linear      PAT 01-DEC-1998
DEFINITION Sequence 9 from patent US 5728518.
ACCESSION  I92348
VERSION     I92348.1 GI:3936818
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 81)
AUTHORS   Carmichael,E.
TITLE     Antiviral poly- and oligonucleotides
JOURNAL   Patent: US 5728518-A 9 17-MAR-1998;
FEATURES   Location/Qualifiers
            source
            1..81
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 81;
Best Local Similarity 95.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAAGGUCGA 20
    |||||
Db 81 AAATCTTTTATAAGGTCGA 62

RESULT 6
AY329529/c
LOCUS       AY329529             253 bp     DNA      linear      VRL 08-JUN-2004
DEFINITION Hepatitis B virus isolate A611252E X protein gene, partial cds; and
           prec/C protein gene, complete cds.
ACCESSION  AY329529
VERSION     AY329529.1 GI:37625315
KEYWORDS   .
SOURCE     Hepatitis B virus
ORGANISM   Hepatitis B virus
           Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE  1 (bases 1 to 253)
AUTHORS   Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
           Silva,L.C. and Carrilho,F.J.
TITLE     Hepatitis B Virus Genotypes and Precore and Core Mutants in
           Brazilian Patients
JOURNAL   J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
PUBMED    15184419
REFERENCE  2 (bases 1 to 253)
AUTHORS   Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
           Bernardini,A.P.

```

```

TITLE      Direct Submission
JOURNAL    Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES   Location/Qualifiers
            source
            1..253
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /isolate="A611252B"
            /db_xref="taxon:10407"
            <1..158
            /codon_start=3
            /product="X protein"
            /protein_id="AAQ95861.1"
            /db_xref="GI:37625316"
            /translation="STTDLEAYFKDCLFKDWELGEBETRLMIFVLGGCRHKLVCAPAS
            CNFFTSA"
            134..217
            /codon_start=1
            /product="prec/C protein"
            /protein_id="AAQ95862.1"
            /db_xref="GI:37625317"
            /translation="MQLFHLCLIVISCTCTFQASKLCIGWL"
            CNFFTSA"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 95.0%; Pred. No. 89;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAATCTTTTATAAGGUGCA 20
        |||||
DB      247 AAATCTTTTATAAGGTCGA 228
        |||||

RESULT 7
AY329561/c
LOCUS      Hepatitis B virus isolate D272811E X protein gene, partial cds; and
DEFINITION prec/C protein gene, complete cds.
ACCESSION  AY329561
VERSION     AY329561.1 GI:37625410
KEYWORDS
SOURCE      Hepatitis B virus
ORGANISM    Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1 (bases 1 to 253)
AUTHORS     Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
TITLE       Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
JOURNAL     J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
PUBMED     15184419
REFERENCE   2 (bases 1 to 253)
AUTHORS     Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
TITLE       Direct Submission
JOURNAL     Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES    Location/Qualifiers
            source
            1..253
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /isolate="D273984E"
            /db_xref="taxon:10407"
            <1..158
            /codon_start=3
            /product="X protein"
            /protein_id="AAQ95926.1"
            /db_xref="GI:37625414"
            /translation="STTDLEAYFKDCLFKDWELGEBEIRLMIFVLGGCRHKLVCAPAP
            CNFFTSA"
            134..217
            /codon_start=1
            /product="prec/C protein"
            /protein_id="AAQ95927.1"
            /db_xref="GI:37625415"
            /translation="MQLFHLCLIIISCSCTPTQASKLCIGWL"
            CNFFTSA"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 95.0%; Pred. No. 89;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAATCTTTTATAAGGUGCA 20
        |||||
DB      247 AAATCTTTTATAAGGTCGA 228
        |||||

RESULT 8
AY329562/c
LOCUS      Hepatitis B virus isolate D273984E X protein gene, partial cds; and
DEFINITION prec/C protein gene, complete cds.
ACCESSION  AY329562
VERSION     AY329562.1 GI:37625413
KEYWORDS
SOURCE      Hepatitis B virus
ORGANISM    Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1 (bases 1 to 253)
AUTHORS     Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
TITLE       Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
JOURNAL     J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
PUBMED     15184419
REFERENCE   2 (bases 1 to 253)
AUTHORS     Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
TITLE       Direct Submission
JOURNAL     Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES    Location/Qualifiers
            source
            1..253
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /isolate="D273984E"
            /db_xref="taxon:10407"
            <1..158
            /codon_start=3
            /product="X protein"
            /protein_id="AAQ95926.1"
            /db_xref="GI:37625414"
            /translation="STTDLEAYFKDCLFKDWELGEBEIRLMIFVLGGCRHKLVCAPAP
            CNFFTSA"
            134..217
            /codon_start=1
            /product="prec/C protein"
            /protein_id="AAQ95927.1"
            /db_xref="GI:37625415"
            /translation="MQLFHLCLIIISCSCTPTQASKLCIGWL"
            CNFFTSA"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 95.0%; Pred. No. 89;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAATCTTTTATAAGGUGCA 20
        |||||
DB      247 AAATCTTTTATAAGGTCGA 228
        |||||

RESULT 9
AY329568/c
LOCUS      Hepatitis B virus isolate D272811E X protein gene, partial cds; and
DEFINITION prec/C protein gene, complete cds.
ACCESSION  AY329561
VERSION     AY329561.1 GI:37625410
KEYWORDS
SOURCE      Hepatitis B virus
ORGANISM    Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE   1 (bases 1 to 253)
AUTHORS     Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
            Silva,L.C. and Carrilho,F.J.
TITLE       Hepatitis B Virus Genotypes and Precore and Core Mutants in
            Brazilian Patients
JOURNAL     J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
PUBMED     15184419
REFERENCE   2 (bases 1 to 253)
AUTHORS     Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and
            Bernardini,A.P.
TITLE       Direct Submission
JOURNAL     Submitted (23-JUN-2003) Research & Development, Laboratorio
            Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
            01402-001, Brazil
FEATURES    Location/Qualifiers
            source
            1..253
            /organism="Hepatitis B virus"
            /mol_type="genomic DNA"
            /isolate="D272811E"
            /db_xref="taxon:10407"
            <1..158
            /codon_start=3
            /product="X protein"
            /protein_id="AAQ95924.1"
            /db_xref="GI:37625411"
            /translation="STTDLEAYFKDCLFKDWELGEBETRLMIFVLGGCRHKLVCAPAP
            CNFFTSA"
            134..217
            /codon_start=1
            /product="prec/C protein"
            /protein_id="AAQ95925.1"
            /db_xref="GI:37625412"
            /translation="MQLFHLCLIIISCSCTPTQASKLCIGWL"
            CNFFTSA"

ORIGIN
Query Match      100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 95.0%; Pred. No. 89;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAATCTTTTATAAGGUGCA 20
        |||||
DB      247 AAATCTTTTATAAGGTCGA 228
        |||||
```

LOCUS	AY329568	253 bp	DNA	linear	VRL 08-JUN-2004			
DEFINITION	Hepatitis B virus isolate D296668E X protein gene, partial cds; and prec/C protein gene, complete cds.							
ACCESSION	AY329568							
VERSION	AY329568.1	GI:37625431						
KEYWORDS								
SOURCE	Hepatitis B virus							
ORGANISM	Hepatitis B virus							
REFERENCE	1 (bases 1 to 253)							
AUTHORS	Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da Silva,L.C. and Carrilho,F.J.							
TITLE	Hepatitis B Virus Genotypes and Precore and Core Mutants in Brazilian Patients							
JOURNAL	J. Clin. Microbiol. 42 (6), 2455-2460 (2004)							
PUBMED	15184419							
REFERENCE	2 (bases 1 to 253)							
AUTHORS	Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and Bernardini,A.P.							
TITLE	Direct Submission							
JOURNAL	Submitted (23-JUN-2003) Research & Development, Laboratorio Bioquimico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo 01402-001, Brazil							
FEATURES	Location/Qualifiers							
source	1..253							
CDS	/organism="Hepatitis B virus"							
	/mol_type="genomic DNA"							
	/isolate="D296668E"							
	/db_xref="taxon:10407"							
	<1..158							
CDS	/codon_start=3							
	/product="X protein"							
	/protein_id="AA095938.1"							
	/db_xref="GI:37625432"							
	/translation="STTDLEAYFKDCLFQDWELGBELRLMIFVLGGCRHLVCAPAPCNFTTSA"							
CDS	134..217							
	/codon_start=1							
	/product="prec/C protein"							
	/protein_id="AA095939.1"							
	/db_xref="GI:37625433"							
ORIGIN	/translation="MQLFHLCLLISCSPTVQASKLGLWL"							
Query Match	100.0%; Score 20; DB 14; Length 253;							
Best Local Similarity	95.0%; Pred. No. 89;							
Matches	19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;							
Qy	1 AAATTCTTTTATAGGCGUCA 20							
Db								
	247 AAATTCTTTTATAGGCGTCA 228							
RESULT 10								
AY329573/c								
LOCUS	AY329573	253 bp	DNA	linear	VRL 08-JUN-2004			
DEFINITION	Hepatitis B virus isolate D604917E X protein gene, partial cds; and prec/C protein gene, complete cds.							
ACCESSION	AY329573							
VERSION	AY329573.1	GI:37625446						
KEYWORDS								
SOURCE	Hepatitis B virus							
ORGANISM	Hepatitis B virus							
REFERENCE	1 (bases 1 to 253)							
AUTHORS	Sitnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da Silva,L.C. and Carrilho,F.J.							
TITLE	Hepatitis B Virus Genotypes and Precore and Core Mutants in Brazilian Patients							
JOURNAL	J. Clin. Microbiol. 42 (6), 2455-2460 (2004)							
PUBMED	15184419							
REFERENCE	2 (bases 1 to 253)							
AUTHORS	Rebello Pinho,J.R., Sitnik,R., Carrilho,F.J., Da Silva,L.C. and Bernardini,A.P.							

```

CDS
134. .217
/codon_start=1
/product="preC/C protein"
/protein_id="AA095953.1"
/db_xref="GI:37625454"
/translation="MQLFHLCLIISCSCTPTVQASKLCLGWL"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 95.0%; Pred. No. 89;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATCTTTATAGGUGCA 20
|||||
Db 247 AAATCTTTATAGGTCGA 228

RESULT 12
AY329581/c
LOCUS
DEFINITION
Hepatitis B virus isolate D639472E X protein gene, partial cds; and
prec/C protein gene, complete cds.
ACCESSION
AY329581
VERSION
AY329581.1 GI:37625470
KEYWORDS
Hepatitis B virus
SOURCE
Hepatitis B virus
ORGANISM
Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 253)
Sítnik,R., Rebello Pinho,J.R., Bertolini,D.A., Bernardini,A.P., Da
Silva,L.C. and Carrilho,F.J.
Hepatitis B Virus Genotypes and Precore and Core Mutants in
Brazilian Patients
J. Clin. Microbiol. 42 (6), 2455-2460 (2004)
PUBMED
15184419
REFERENCE
2 (bases 1 to 253)
Rebello Pinho,J.R., Sítnik,R., Carrilho,F.J., Da Silva,L.C. and
Bernardini,A.P.
Direct Submission
Submitted (23-JUN-2003) Research & Development, Laboratório
Bioquímico Jardim Paulista, Av Brig Luiz Antonio 4701, Sao Paulo
01402-001, Brazil
Location/Qualifiers
1. .253
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D639472E"
/db_xref="taxon:10407"
<1..158
/codon_start=3
/product="X protein"
/protein_id="AA095954.1"
/db_xref="GI:37625471"
/translations="STIDLEAYPKCLFKDWEELGELRLLIPLVGGCRHKLVCAPAP
CNFFTSA"
134. .217
/codon_start=1
/product="preC/C protein"
/protein_id="AA095965.1"
/db_xref="GI:37625472"
/translation="MQLFHLCLIISCSCTPTVQASKLCLGWL"

ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 253;
Best Local Similarity 95.0%; Pred. No. 89;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATCTTTATAGGUGCA 20
|||||
Db 247 AAATCTTTATAGGTCGA 228

RESULT 13

```

```

AF390000/c
LOCUS
DEFINITION
Hepatitis B virus isolate D3 X protein gene, partial cds; and
nonfunctional precore/core protein gene, partial sequence.
ACCESSION
AF390000
VERSION
AF390000.1 GI:16266099
KEYWORDS
Hepatitis B virus
SOURCE
Hepatitis B virus
ORGANISM
Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (bases 1 to 294)
Castro,L.D., Niel,C. and Gomes,S.A.
Low frequency of mutations in the core promoter and precore regions
of hepatitis B virus in anti-HBe positive Brazilian carriers
BMC Microbiol. 1 (1), 10 (2001)
PUBMED
11472634
REFERENCE
2 (bases 1 to 294)
De Castro,L., Niel,C. and Gomes,S.A.
Direct Submission
Submitted (11-JUN-2001) Virology, FIOCRUZ, Av. Brasil 4365, Rio de
Janeiro, RJ 21045-900, Brazil
Location/Qualifiers
1. .294
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/isolate="D3"
/db_xref="taxon:10407"
<1..119
/codon_start=3
/product="X protein"
/protein_id="AA16752.1"
/db_xref="GI:16266100"
/translation="FKDWEELGDSRLMIYVLGGCRHKLVCAPAPCNFFTSA"
95. .>294
/notes="nonfunctional precore/core protein due to mutation"

misc_feature
ORIGIN
Query Match 100.0%; Score 20; DB 14; Length 294;
Best Local Similarity 95.0%; Pred. No. 87;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAATCTTTATAGGUGCA 20
|||||
Db 208 AAATCTTTATAGGTCGA 189

RESULT 14
HPBHBD/c
LOCUS
DEFINITION
Hepatitis B virus precore and core protein gene, 5' end of cds.
ACCESSION
L12359
VERSION
L12359.1 GI:306267
KEYWORDS
HBcAg protein; HBeAg protein; core protein; nucleotide binding
protein; precore protein.
SOURCE
Hepatitis B virus
ORGANISM
Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE
1 (sites)
Tong,S.P., Li,J.S., Vitvitski,L. and Trepo,C.
Active hepatitis B virus replication in the presence of anti-HBe is
associated with viral variants containing an inactive pre-C region
Virology 176 (2), 596-603 (1990)
PUBMED
2345966
REFERENCE
2 (bases 1 to 333)
Li,J.S., Tong,S.P., Wen,Y.M., Vitvitski,L., Zhang,Q. and Trepo,C.
Hepatitis B virus genotype A rarely circulates as an HBe-minus
mutant: possible contribution of a single nucleotide in the precore
region
J. Virol. 67 (9), 5402-5410 (1993)
PUBMED
93353617
JOURNAL
MEDLINE
8350403
COMMENT
Original source text: Hepatitis B virus DNA.

```

FEATURES
source

Location/Qualifiers
1. .333
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:10407"
/notes="genotype D; from French HBeAg- patient"
1. .>333
/standard_name="HBeAg"
/pseudo
/codon_start=1
/product="precore protein"
82. .84
/notes="HBe-abolishing"
88. .>333
/standard_name="HBeAb"
/codon_start=1
/protein_id="AA03102.1"
/db_xref="GI:306268"
/translations="MDIDPKYKFGATVELLFLPSPDFPVRDLDTAAALYREALS
PECHSHHTALRQAILCWGELMTLATWVGANLDDPASR"

CDS

variation

CDS

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 333;
Best Local Similarity 95.0%; Pred. No. 85;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATTCCTTTATTAAGGUGCA 20

Db 114 AAATTCCTTTATTAAGGUGCA 95

RESULT 15

AB167603/c 398 bp DNA linear VRL 01-OCT-2004
LOCUS
DEFINITION Hepatitis B virus gene for polyprotein, partial cds, clone: NEP75.
ACCESSION AB167603
VERSION AB167603.1 GI:53148166

KEYWORDS

SOURCE

Hepatitis B virus
Hepatitis B virus
Viruses; Retroviridae; Hepadnaviridae; Orthohepadnavirus.

1 Tanaka, Y., Hasegawa, I., Kato, T., Orito, E. and Mizokami, M.

A Case-control Study for Differences among Hepatitis B Virus

Infections of Genotypes A (Subtypes Aa and Ae) and D

Unpublished

2 (bases 1 to 398)

Tanaka, Y. and Mizokami, M.

Direct Submission

Submitted (15-MAR-2004) Yasuhiro Tanaka, Nagoya City University

Graduate School of Medical Sciences, Department of Clinical

Molecular Informative Medicine; 1 Kawasaki, Mizuho-cho, Mizuho-ku,

Nagoya, Aichi 467-8601, Japan (E-mail: ytanaka@med.nagoya-cu.ac.jp,

Tel: 81-52-853-8292, Fax: 81-52-842-0021)

FEATURES

source

Location/Qualifiers
1. .398
/organism="Hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:10407"
/clone="NEP75"
184. .>398
/notes="precore-core region"
/codon_start=1
/product="polyprotein"
/protein_id="BAD52175.1"
/db_xref="GI:53148167"
/translation="MQLFHLCLISCTPQASKLCGLWGMMDIDPKYKFGATVEL
LSFLPSDFPVRDLDTASALYREAL"

CDS

ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 398;
Best Local Similarity 95.0%; Pred. No. 82;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATTCCTTTATTAAGGUGCA 20
|||||
Db 297 AAATTCCTTTATTAAGGUGCA 278
|||||

Search completed: March 17, 2005, 08:14:18
Job time : 683.733 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 04:03:33 ; Search time 171.333 Seconds
(without alignments)
691.020 Million cell updates/sec

Title: US-08-901-612A-65
Perfect score: 20
Sequence: 1 aaattctttataaggguaga 20
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	2	AAT72579 Hepatitis
2	20	100.0	20	2	AAT72580 Hepatitis
3	20	100.0	27	2	AAT72577 Hepatitis
C 4	20	100.0	28	2	AAQ80499 Primer to
C 5	20	100.0	28	2	Aav45201 Primer MF
C 6	20	100.0	28	3	AAa62585 Transgeni
C 7	20	100.0	28	12	Adn36074 Probe #15
C 8	20	100.0	28	12	Adn36080 Probe #16
C 9	20	100.0	28	12	Adn36073 Probe #15
C 10	20	100.0	28	12	Adn36075 Probe #15
C 11	20	100.0	28	12	Adn36078 Probe #15
C 12	20	100.0	28	12	Adn36077 Probe #15
C 13	20	100.0	28	12	Adn36076 Probe #15
C 14	20	100.0	28	12	Adn36072 Probe #16
C 15	20	100.0	28	12	Adn36079 Probe #16
C 16	20	100.0	35	2	AAx36590 PCR prime
C 17	20	100.0	35	8	Abx95880 PCR prime
C 18	20	100.0	35	10	Ac07807 Hepatitis
C 19	20	100.0	40	10	Adel0983 Chimeric
C 20	20	100.0	40	10	Adel10985 Chimeric

C 21	20	100.0	40	12	ADG64056
C 22	20	100.0	40	12	ADG64054
C 23	20	100.0	40	12	ADP73665
C 24	20	100.0	40	12	ADP73667
C 25	20	100.0	40	13	ADR12912
C 26	20	100.0	40	13	ADR12910
C 27	20	100.0	41	10	ADG46976
C 28	20	100.0	41	11	ADM83221
C 29	20	100.0	52	10	ADE11060
C 30	20	100.0	52	10	ADG47004
C 31	20	100.0	52	11	ADM83249
C 32	20	100.0	52	12	ADG64131
C 33	20	100.0	52	12	ADP73784
C 34	20	100.0	52	13	ADR12987
C 35	20	100.0	53	12	ADN36055
C 36	20	100.0	59	10	ADE11058
C 37	20	100.0	59	10	ADG47002
C 38	20	100.0	59	11	ADM83247
C 39	20	100.0	59	12	ADG64129
C 40	20	100.0	59	12	ADP73782
C 41	20	100.0	59	13	ADR12985
C 42	20	100.0	504	11	ADM41005
C 43	20	100.0	504	11	ADM41004
C 44	20	100.0	513	6	ABK67524
C 45	20	100.0	513	6	ABK67525

ALIGNMENTS

RESULT 1
AAT72579
ID AAT72579 standard; DNA; 20 BP.
XX
AC AAT72579;
XX
DT 04-SEP-1997 (first entry)
XX
DE Hepatitis B virus RNA antisense oligonucleotide HBV93b.
KW HBV; HBV infection; inhibition; replication; ss.
XX
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..20
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"

XX
PN W09639502-A1.
XX
PD 12-DEC-1996.
XX
PF 04-JUN-1996; 96WO-EP002432.
PR 06-JUN-1995; 95US-00467397.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (HYBR-) HYBRIDON INC.
XX
PI Craig CJ, Frank BL, Goodchild J, Jupp R, Kiluskie RE, Mills JS;
PI Roberts NA, Roberts PC, Slade A;
DR WPI; 1997-043124/04.
XX
PT Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
PT used in the detection and treatment of HBV infection.
XX
PS Claim 1; Page 12; 81pp; English.
XX
CC The present sequence represents a synthetic oligonucleotide HBV93b which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a

Adg64056	Recombina
Adg64054	Recombina
Adp73665	HBV pKK-2
Adp73667	HBV pKK-2
Adr12912	HBV Hbc-C
Adr12910	HBV Hbc-C
Adg46976	PCR prime
Adm83221	PCR prime
Adel1060	Chimeric
Adg47004	PCR prime
Adm83249	Influenza
Adg64131	Recombina
Adp73784	Influenza
Adr12987	Influenza
Adn36055	Probe #13
Adel1058	Chimeric
Adg47002	PCR prime
Adm83247	Influenza
Adg64129	Recombina
Adp73782	Influenza
Adr12985	Influenza
Adm41005	Hbc relat
Adm41004	Hbc relat
Abk67524	DNA encod
Abk67525	DNA encod

CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non- contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection

XX Sequence 20 BP; 7 A; 2 C; 4 G; 7 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 95.0%; Pred. No. 7;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAAGGUCGA 20
DB 1 AAATCTTTTATAAGGTCGA 20
|||||

RESULT 2
AAT72580
ID AAT72580 standard; DNA; 20 BP.

XX AAT72580;
AC
XX 04-SEP-1997 (first entry)
DT Hepatitis B virus RNA antisense oligonucleotide HBV93MB.
DE HBV; HBV infection; inhibition; replication; ss.
XX Synthetic.

XX Key Location/Qualifiers
FH misc_feature 1..20
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"
FT misc_RNA 16..20
FT /tag= b
FT /note= "2'-OMe RNA"
FT modified_base 16
FT /tag= c
FT /mod_base= gm
FT modified_base 17
FT /tag= d
FT /mod_base= um
FT modified_base 18
FT /tag= e
FT /mod_base= cm
FT modified_base 19
FT /tag= f
FT /mod_base= gm
FT modified_base 20
FT /tag= g
FT /mod_base= OTHER
FT /note= "2'-O-methyladenosine"

XX WO9639502-A1.
XX
XX 12-DEC-1996.
XX
XX 04-JUN-1996; 96WO-EP002432.
XX
XX 06-JUN-1995; 95US-00467397.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX (HYBR-) HYBRIDON INC.
XX
XX Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
XX Roberts NA, Roberts PC, Slade A;
XX WPI; 1997-043124/04.
XX
XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
XX used in the detection and treatment of HBV infection.

PS Claim 1; Page 12; 81pp; English.

XX The present sequence represents a synthetic oligonucleotide HBV93MB which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a
CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non- contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection

XX Sequence 20 BP; 7 A; 2 C; 4 G; 6 T; 1 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAAGGUCGA 20
DB 1 AAATCTTTTATAAGGUCGA 20
|||||

RESULT 3
AAT72577
ID AAT72577 standard; DNA; 27 BP.

XX AAT72577;
AC
XX 04-SEP-1997 (first entry)
DT Hepatitis B virus RNA antisense oligonucleotide HBV94b.
DE HBV; HBV infection; inhibition; replication; ss.
XX Synthetic.

XX Key Location/Qualifiers
FH misc_feature 1..27
FT /tag= a
FT /note= "Internucleotide linkages are phosphorothioate"

XX WO9639502-A1.

XX 12-DEC-1996.

XX 04-JUN-1996; 96WO-EP002432.

XX 06-JUN-1995; 95US-00467397.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX (HYBR-) HYBRIDON INC.

XX Craig CJ, Frank BL, Goodchild J, Jupp R, Kilkuskie RE, Mills JS;
XX Roberts NA, Roberts PC, Slade A;

XX WPI; 1997-043124/04.

XX Oligo:nucleotide(s) complementary to hepatitis B virus (HBV) sequences -
XX used in the detection and treatment of HBV infection.

XX Claim 1; Page 12; 81pp; English.

XX The present sequence represents a synthetic oligonucleotide HBV94b which
CC is complementary to a portion of the hepatitis B virus (HBV) RNA. The
CC antisense oligonucleotide may be used to detect the presence of HBV in a
CC sample. The antisense oligonucleotide, and oligonucleotides containing a
CC sequence which is complementary to at least two non- contiguous regions
CC of an HBV nucleic acid, may be used for inhibiting HBV replication in a
CC cell or for the treatment of HBV infection

XX Sequence 27 BP; 8 A; 4 C; 5 G; 10 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 27;
Best Local Similarity 95.0%; Pred. No. 7.1;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAGGCGCA 20
 |||||:|||||
 Db 1 AAATCTTTTATAGGCGCA 20

RESULT 4
 AAQ80499/C
 ID AAQ80499 standard; DNA; 28 BP.
 XX
 AC AAQ80499;
 XX
 XX 25-MAR-2003 (revised)
 DT 23-AUG-1995 (first entry)
 XX
 XX Primer to amplify hepatitis B virus core region.
 XX
 XX hepatitis B virus; X region; core region; primer; PCR; amplification;
 KW polymerase chain reaction; detection; viral infection; ss.
 XX
 OS Synthetic.
 XX
 XX WO9429483-A1.
 XX
 XX 22-DEC-1994.
 XX
 XX 03-JUN-1994; 94WO-US006360.
 XX
 XX 08-JUN-1993; 93US-00074346.
 XX
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 PA
 XX Feitelson M, Duan L, Guo J;
 PI
 XX WPI; 1995-036505/05.
 DR
 XX Detection of hepatitis B virus (HBV) variants having deletions in the X
 PT region - by detection of antibodies against HBV polymerase and HB X
 PT antigen.
 XX
 PS Claim 3; Page 34; 45pp; English.
 XX
 CC This primer designated MF03 covers nucleotide bases 1903-1949 at the
 CC beginning of the hepatitis B virus (HBV) core open reading frame. It is
 CC used with MF04 (AAQ80500) to amplify the core gene. The primers allow the
 CC detection of a specific class of HBV variants. They are useful for
 CC demonstrating the presence of productive virus infection and may prove
 CC useful in monitoring therapeutics. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX
 SQ Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 28;
 Best Local Similarity 95.0%; Pred. No. 7.1;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAGGCGCA 20
 |||||:|||||
 Db 27 AAATCTTTTATAGGCGCA 8

RESULT 5
 AAV45201/C
 ID AAV45201 standard; DNA; 28 BP.
 XX
 AC AAV45201;
 XX
 XX 19-OCT-1998 (first entry)
 DT
 XX
 DE Primer MF03.
 XX
 KW ss; PCR; primer; amplification; viral infection; bacterial infection;

immune response; hepatitis B virus.
 Mus sp.
 WO9829121-A1.
 09-JUL-1998.
 02-JAN-1998; 98WO-US004116.
 02-JAN-1997; 97US-0034596P.
 (UYJE-) UNIV JEFFERSON THOMAS.
 Michaels F, Block T;
 WPI; 1998-387782/33.
 Modulating immune responses in mammals infected with infectious agent(s)
 - e.g. to reduce pathogenicity caused by immune responses in cases where
 the infectious agent has limited pathogenicity.
 Example 2; Page 37; 55pp; English.
 The primers AAV45201 and AAV45202 were used to detect the presence of a
 HBV genome which had been microinjected into embryos of SCID mice in an
 example to demonstrate modulating an immune response in a mammal infected
 with an infectious agent. This comprises transcutaneous administration of a
 composition comprising an epitope which is located in close proximity to
 the immune response. The process may be used in treatment of mammals
 which are acutely or chronically infected with infectious agents, such as
 viruses or bacteria. It may be used to increase the immune response, or
 it may be used to decrease the immune response in cases where the
 infectious agent itself exhibits limited pathogenicity but the immune
 response to the infectious agent causes more significant pathogenicity.
 This can be the case in, e.g. hepatitis B virus (HBV) infection. The
 process can modulate undesirable autoimmune responses exhibited by
 mammals infected with viral, bacterial and parasitic agents. It can
 prevent life-long disabilities which result from these infections
 Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 28;
 Best Local Similarity 95.0%; Pred. No. 7.1;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAGGCGCA 20
 |||||:|||||
 Db 27 AAATCTTTTATAGGCGCA 8

RESULT 6
 AA62585/C
 ID AA62585 standard; DNA; 28 BP.
 XX
 AC AA62585;
 XX
 XX 22-NOV-2000 (first entry)
 DT
 XX
 DE Transgenic SCID mouse hepatitis virus transgene PCR primer MF03.
 XX
 KW Mouse; SCID; severe combined immunodeficiency; transgenic mouse;
 KW hepatitis virus; hepatitis B; hepatitis C; chronic liver disease;
 KW PCR primer; ss.
 XX
 OS Mus sp.
 XX
 XX US6087556-A.
 PN
 XX 11-JUL-2000.
 PD
 XX 07-JAN-1998; 98US-00003200.
 PF
 XX

CC 604) in the DNA polymerase region of HBV. The method comprises
 CC determining whether the HBV carried by the host bears one or more of the
 CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
 CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
 CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
 CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
 CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
 CC and C1962A representing specific double point mutations in the DNA
 CC precore/core promoter or ORF region. The method and kit is useful in
 CC predicting the long-term response of a host of HBV to 3TC therapy (also
 CC known as lamivudine). This sequence represents an oligonucleotide
 CC sequence used in the method of the invention to detect a mutation in the
 CC above mentioned sequences.

XX
 SQ Sequence 28 BP; 8 A; 6 C; 5 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
 Best Local Similarity 95.0%; Pred. No. 7.1;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTCTTATAAGGGGUCGA 20
 |||||
 Db 20 AAATCTCTTATAAGGGGTCGA 1

RESULT 9
 ADN36073/C
 ID ADN36073 standard; DNA; 28 BP.
 XX
 AC ADN36073;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Probe #154 to determine effect of long term lamivudine treatment of HBV.
 XX
 KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
 XX
 OS Hepatitis B virus.
 XX
 PN WO2004031729-A2.
 XX
 PD 15-APR-2004.
 XX
 PF 01-OCT-2003; 2003WO-US031121.
 XX
 PR 01-OCT-2002; 2002US-0415301P.
 XX
 PA (GEOU) UNIV GEORGETOWN.
 XX
 PI Korba BE, Cincio A, Gerin JL;
 XX
 DR WPI; 2004-348004/32.
 XX
 PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
 PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
 PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
 XX
 PS Claim 31; SEQ ID NO 154; 107pp; English.

XX The invention relates to a method of predicting the long term response of
 CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
 CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
 CC leucine at amino acid position (aa) 91 in the DNA polymerase region
 CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
 CC 604) in the DNA polymerase region of HBV. The method comprises
 CC determining whether the HBV carried by the host bears one or more of the
 CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
 CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
 CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
 CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
 CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
 CC and C1962A representing specific double point mutations in the DNA
 CC precore/core promoter or ORF region. The method and kit is useful in

CC predicting the long-term response of a host of HBV to 3TC therapy (also
 CC known as lamivudine). This sequence represents an oligonucleotide
 CC sequence used in the method of the invention to detect a mutation in the
 CC above mentioned sequences.

XX
 SQ Sequence 28 BP; 10 A; 5 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
 Best Local Similarity 95.0%; Pred. No. 7.1;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTCTTATAAGGGGUCGA 20
 |||||
 Db 27 AAATCTCTTATAAGGGGTCGA 8

RESULT 10
 ADN36075/C
 ID ADN36075 standard; DNA; 28 BP.
 XX
 AC ADN36075;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Probe #156 to determine effect of long term lamivudine treatment of HBV.
 XX
 KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.
 XX
 OS Hepatitis B virus.
 XX
 PN WO2004031729-A2.
 XX
 PD 15-APR-2004.
 XX
 PF 01-OCT-2003; 2003WO-US031121.
 XX
 PR 01-OCT-2002; 2002US-0415301P.
 XX
 PA (GEOU) UNIV GEORGETOWN.
 XX
 PI Korba BE, Cincio A, Gerin JL;
 XX
 DR WPI; 2004-348004/32.

XX Predicting the long-term response of a host of hepatitis B virus (HBV) to
 PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
 PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
 XX
 PS Claim 31; SEQ ID NO 156; 107pp; English.

XX The invention relates to a method of predicting the long term response of
 CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
 CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
 CC leucine at amino acid position (aa) 91 in the DNA polymerase region
 CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
 CC 604) in the DNA polymerase region of HBV. The method comprises
 CC determining whether the HBV carried by the host bears one or more of the
 CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
 CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
 CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
 CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
 CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
 CC and C1962A representing specific double point mutations in the DNA
 CC precore/core promoter or ORF region. The method and kit is useful in
 CC predicting the long-term response of a host of HBV to 3TC therapy (also
 CC known as lamivudine). This sequence represents an oligonucleotide
 CC sequence used in the method of the invention to detect a mutation in the
 CC above mentioned sequences.

XX
 SQ Sequence 28 BP; 10 A; 5 C; 6 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
 Best Local Similarity 95.0%; Pred. No. 7.1;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAGGGUCGA 20
 |||||
 Db 25 AAATCTTTTATAGGGTCGA 6

RESULT 11

ADN36078/c
 ID ADN36078 standard; DNA; 28 BP.

XX AC ADN36078;

XX XX 01-JUL-2004 (first entry)

XX XX Probe #159 to determine effect of long term lamivudine treatment of HBV.

XX DE ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.

XX KW Hepatitis B virus.

XX OS WO2004031729-A2.

XX PN 15-APR-2004.

XX PD 01-OCT-2003; 2003WO-US031121.

XX PF 01-OCT-2002; 2002US-0415301P.

XX PR (GEOU) UNIV GEORGETOWN.

XX PA Korba BE, Ciano A, Gerin JL;

XX PI WPI; 2004-348004/32.

XX DR Predicting the long-term response of a host of hepatitis B virus (HBV) to
 XX PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
 XX PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.

XX PS Claim 31; SEQ ID NO 159; 107pp; English.

XX CC The invention relates to a method of predicting the long term response of
 XX CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
 XX CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
 XX CC leucine at amino acid position (aa) 91 in the DNA polymerase region
 XX CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
 XX CC 604) in the DNA polymerase region of HBV. The method comprises
 XX CC determining whether the HBV carried by the host bears one or more of the
 XX CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
 XX CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
 XX CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
 XX CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
 XX CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
 XX CC precore/core promoter or ORF region. The method and kit is useful in
 XX CC predicting the long-term response of a host of HBV to 3TC therapy (also
 XX CC known as lamivudine). This sequence represents an oligonucleotide
 XX CC sequence used in the method of the invention to detect a mutation in the
 XX CC above mentioned sequences.

XX SQ Sequence 28 BP; 9 A; 6 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
 Best Local Similarity 95.0%; Pred. No. 7.1;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAGGGUCGA 20
 |||||
 Db 22 AAATCTTTTATAGGGTCGA 3

RESULT 12

ADN36077/c

ID ADN36077 standard; DNA; 28 BP.

XX AC ADN36077;

XX XX 01-JUL-2004 (first entry)

XX DE Probe #158 to determine effect of long term lamivudine treatment of HBV.

XX KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.

XX OS Hepatitis B virus.

XX PN WO2004031729-A2.

XX PD 15-APR-2004.

XX PF 01-OCT-2003; 2003WO-US031121.

XX PR 01-OCT-2002; 2002US-0415301P.

XX PA (GEOU) UNIV GEORGETOWN.

XX PI Korba BE, Ciano A, Gerin JL;

XX DR WPI; 2004-348004/32.

XX PT Predicting the long-term response of a host of hepatitis B virus (HBV) to
 XX PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
 XX PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.

XX PS Claim 31; SEQ ID NO 158; 107pp; English.

XX CC The invention relates to a method of predicting the long term response of
 XX CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
 XX CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
 XX CC leucine at amino acid position (aa) 91 in the DNA polymerase region
 XX CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
 XX CC 604) in the DNA polymerase region of HBV. The method comprises
 XX CC determining whether the HBV carried by the host bears one or more of the
 XX CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
 XX CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
 XX CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
 XX CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
 XX CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
 XX CC precore/core promoter or ORF region. The method and kit is useful in
 XX CC predicting the long-term response of a host of HBV to 3TC therapy (also
 XX CC known as lamivudine). This sequence represents an oligonucleotide
 XX CC sequence used in the method of the invention to detect a mutation in the
 XX CC above mentioned sequences.

XX SQ Sequence 28 BP; 10 A; 6 C; 5 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
 Best Local Similarity 95.0%; Pred. No. 7.1;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTTATAGGGUCGA 20
 |||||
 Db 23 AAATCTTTTATAGGGTCGA 4

RESULT 13

ADN36076/c

ID ADN36076 standard; DNA; 28 BP.

XX AC ADN36076;

XX XX 01-JUL-2004 (first entry)

XX DE Probe #157 to determine effect of long term lamivudine treatment of HBV.

XX KW ss; probe; hepatitis B virus; HBV; 3TC therapy; mutation; lamivudine.

PT 3TC therapy comprises determining whether the HBV bears a nucleic acid
 PT encoding leucine at amino acid position (aa) 91 or cysteine at aa256.
 XX
 PS Claim 31; SEQ ID NO 160; 107pp; English.
 XX
 CC The invention relates to a method of predicting the long term response of
 CC a host of hepatitis B virus (HBV) to 3TC therapy by determining whether
 CC the HBV carried by the host (i) bears a nucleic acid that encodes for a
 CC leucine at amino acid position (aa) 91 in the DNA polymerase region
 CC (originally codon 438) or a (ii) a cysteine at aa256 (originally codon
 CC 604) in the DNA polymerase region of HBV. The method comprises
 CC determining whether the HBV carried by the host bears one or more of the
 CC following mutations: (i) Q213S (glutamine to serine at aa213) (originally
 CC codon 604) in the HBV polymerase region, (ii) G1739T, A1752C/T, T1909C,
 CC T1960G, or T1961A/G specific point mutation in the DNA precore/core
 CC promoter or open reading frame (ORF) region or (iii) a pair of nucleotide
 CC changes A1738C and G1739T, A1750G and A1752G, T1909G and A1911T or T1961A
 CC and C1962A representing specific double point mutations in the DNA
 CC precore/core promoter or ORF region. The method and kit is useful in
 CC predicting the long-term response of a host of HBV to 3TC therapy (also
 CC known as lamivudine). This sequence represents an oligonucleotide
 CC sequence used in the method of the invention to detect a mutation in the
 CC above mentioned sequences.
 XX
 SQ Sequence 28 BP; 9 A; 5 C; 5 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 28;
 Best Local Similarity 95.0%; Pred. No. 7.1;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAATTCTTTATAGGUGCA 20
 DB 21 AAATTCTTTATAGGTCGA 2

Search completed: March 17, 2005, 06:48:45
 Job time : 172.333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2005, 05:44:58 ; Search time 1386.27 Seconds
(without alignments)
549.162 Million cell updates/sec

Title: US-08-901-612a-65

Perfect score: 20

Sequence: 1 aaattcttataaggguca 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	18	90.0	866	9	CL798479
C 2	17.4	87.0	624	6	CA029935
C 3	17.4	87.0	893	7	CK125419
C 4	17.4	87.0	975	9	CNS06GLD
C 5	17.4	87.0	1003	9	CNS06J82
C 6	17.4	87.0	1028	9	AL401256
C 7	16.8	84.0	85	9	AG265169
C 8	16.8	84.0	293	2	BB009906
C 9	16.8	84.0	299	1	AV258784
C 10	16.8	84.0	407	6	CA856765
C 11	16.8	84.0	462	4	B1926889
C 12	16.8	84.0	494	5	BP527047
C 13	16.8	84.0	508	9	CE435953
C 14	16.8	84.0	572	9	CR154100
C 15	16.8	84.0	588	6	CD865385
C 16	16.8	84.0	593	7	CF424535
C 17	16.8	84.0	620	7	CK537618
C 18	16.8	84.0	626	3	AG016570
C 19	16.8	84.0	626	3	AY068711
C 20	16.8	84.0	630	9	AG016569
C 21	16.8	84.0	678	9	AG119858
C 22	16.8	84.0	700	7	CK569388
C 23	16.8	84.0	707	5	BU041523
C 24	16.8	84.0	707	9	CE761259

25	16.8	84.0	750	9	AG597534	Mus muscu
26	16.8	84.0	752	9	CL748062	OR_BBa011
c 27	16.8	84.0	810	9	CC895730	ZMMBB022
28	16.8	84.0	836	9	CL620368	OR_BBa001
29	16.8	84.0	942	9	CL118472	ISB1-70N2
30	16.8	84.0	1001	9	CNS06N7K	AL406422
c 31	16.8	84.0	1131	8	CC240216	T3 end of
32	16.4	82.0	168	8	BH783463	CH261-121
c 33	16.4	82.0	312	9	CL834207	BH783463 f2mb013f0
c 34	16.4	82.0	323	9	CL728817	OR_CBA005
c 35	16.4	82.0	338	9	CL782534	OR_BBa006
c 36	16.4	82.0	350	1	AI556720	OR_BBa009
c 37	16.4	82.0	431	9	CL611235	UI-R-C2P-
38	16.4	82.0	467	4	BJ267319	OR_BBa000
c 39	16.4	82.0	474	9	CL730394	OR_BBa006
40	16.4	82.0	554	9	CL551502	OB_Ba009
41	16.4	82.0	588	4	BJ271762	BJ271762
42	16.4	82.0	612	5	BM030883	BM030883
43	16.4	82.0	621	2	BB646609	BB646609
c 44	16.4	82.0	637	4	BJ686765	BJ686765
c 45	16.4	82.0	641	4	BJ702377	BJ702377

ALIGNMENTS

RESULT 1
CL798479/c
LOCUS
DEFINITION
OR_CBA0009F12.f OR_CBA Oryza rufipogon genomic clone OR_CBA0009F12
5', Genomic survey sequence.

ACCESSION
CL798479
VERSION
CL798479.1

KEYWORDS
GSS.

SOURCE
Oryza rufipogon

ORGANISM
Oryza rufipogon

REFERENCE
Kim H., Yu Y., Wissotski M., Yost D., Stum D., Rao K., Luo M.,
Jetty R., Kudrna D., Muller C., Hatfield J., Soderlund C. and
Wing R.
Unpublished (2004)

TITLE
OMAP project

JOURNAL
Contact: Rod A. Wing

COMMENT
Arizona Genomics Institute

University of Arizona

Forbes Building Room 303, Tucson, AZ 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GGC ACC CCA

Plate: 0009 row: F column: 12

Seq primer: TAA TAC GAC TCA TAG GG

Class: BAC ends.

Location/Qualifiers

1. .866

/organism="Oryza rufipogon"

/mol_type="genomic DNA"

/db_xref="taxon:4529"

/clone="OR_CBA0009F12"

/tissue_type="young leaves"

/dev_stage="2 week old seedlings"

/lab_host="DH10B T1 phage resistant"

/clone_lib="OR_CBA"

/notes="Vector: pGIBAC1; Site 1: HindIII; Site 2: HindIII; drk treated 36 hrs before harvest"

ORIGIN

Query Match

90.0% Score 18 DB 9 Length 866

```

Best Local Similarity 94.4%; Pred. No. 4.5e+02;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGGUC 18
Db 811 AAATCTTTATAGGGTC 794
|||||

RESULT 2
CA029935/c
LOCUS CA029935 624 bp mRNA linear EST 24-OCT-2002
DEFINITION HX05023r HX Hordeum vulgare subsp. vulgare cDNA clone HX05023
5-PRIME, mRNA sequence.
ACCESSION CA029935.1 GI:243325281
VERSION CA029935
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 624)
AUTHORS Zhang, H., Weschke, W., Michalek, W., Stein, N. and Graner, A.
TITLE EST sequencing and analysis in barley (2002)
JOURNAL Unpublished (2002)
COMMENT Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 5 row: J column: 23
Seq primer: M13rev.

FEATURES
source
Location/Qualifiers
1..624
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:270608"
/db_xref="taxon:112509"
/clone="HX05023"
/tissue_type="apex (3-5 mm in size)"
/dev_stage="X110-Gold"
/lab_host="HX"
/clone_lib="HX"
/notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also recombinants is not 100% reliable."

ORIGIN
Query Match 87.0%; Score 17.4; DB 6; Length 624;
Best Local Similarity 89.5%; Pred. No. 8.5e+02;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGGUC 19
Db 473 AAATCTTCATAGGGTCG 455
|||||

RESULT 3
CK125419/c
LOCUS CK125419 893 bp mRNA linear EST 01-MAR-2004
DEFINITION BES1824107p07 BES1824 Hordeum vulgare subsp. vulgare cDNA clone
MPMG2010P077 5-PRIME, mRNA sequence.
ACCESSION CK125419

Best Local Similarity 94.4%; Pred. No. 4.5e+02;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGGUC 18
Db 811 AAATCTTTATAGGGTC 794
|||||

RESULT 2
CA029935/c
LOCUS CA029935 624 bp mRNA linear EST 24-OCT-2002
DEFINITION HX05023r HX Hordeum vulgare subsp. vulgare cDNA clone HX05023
5-PRIME, mRNA sequence.
ACCESSION CA029935.1 GI:243325281
VERSION CA029935
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 624)
AUTHORS Zhang, H., Weschke, W., Michalek, W., Stein, N. and Graner, A.
TITLE EST sequencing and analysis in barley (2002)
JOURNAL Unpublished (2002)
COMMENT Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 5 row: J column: 23
Seq primer: M13rev.

FEATURES
source
Location/Qualifiers
1..624
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:270608"
/db_xref="taxon:112509"
/clone="HX05023"
/tissue_type="apex (3-5 mm in size)"
/dev_stage="X110-Gold"
/lab_host="HX"
/clone_lib="HX"
/notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also recombinants is not 100% reliable."

ORIGIN
Query Match 87.0%; Score 17.4; DB 6; Length 624;
Best Local Similarity 89.5%; Pred. No. 8.5e+02;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGGUC 19
Db 473 AAATCTTCATAGGGTCG 455
|||||

RESULT 3
CK125419/c
LOCUS CK125419 893 bp DNA linear GSS 30-NOV-2001
DEFINITION T3 end of clone AS0AA005E07 of library AS0AA from strain CLIB 533
of Saccharomyces bayanus, genomic survey sequence.
ACCESSION CK125419
VERSION AL397847
KEYWORDS GSS.

CK125419.1 GI:44808421
EST.
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Pooideae; Triticeae; Hordeum.
1 (bases 1 to 893)
Kramer, A., Feilner, T., Posseling, A., Radchuk, V., Weschke, W.,
Buerkle, L. and Kersten, B.
Application of the protein microarray technology for the
identification of expression library derived target proteins for
barley protein kinase CK2
Unpublished (2003)
Contact: Birgit Kersten* and Winfriede Weschke**
*Plant Protein Chip Group, Department Lehrach, **Department
Molecular Genetics, Gene Expression Group
*Max-Planck-Institute for Molecular Genetics, **Institute of Plant
Genetics and Crop Plant Research Gatersleben
*Innestr. 73, D-14195 Berlin, Germany, **Corrensstrasse 3, D-06466
Gatersleben, Germany
Tel: **49(0)30/84131648, **49(0)394825500
Fax: **49(0)30/84131128, **49(0)394825237
Email: *kersten@molgen.mpg.de, **weschke@ipk-gatersleben.de
Insert Length: 893 Std Error: 0.00
Plate: 7 row: P column: 7
Seq primer: pQR65.
Location/Qualifiers
1..893
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:944977"
/db_xref="taxon:112509"
/clone="MPMGp2010P077"
/tissue_type="embryosac"
/dev_stage="0-10 DAF (days after flowering)"
/lab_host="E. coli, SCS-1/pSE111"
/clone_lib="BES1824"
/notes="Vector: pQR30NST (AF074376); Site 1: SalI; Site 2:
NotI; 0-10 DAF (days after flowering), cDNA synthesis
using pBluescript II XR cDNA-library construction kit
(Stratagen) with an oligo(dT)-primer containing NotI
restriction site and a SalI adapter (Invitrogen). The main
library of 21500 clones was rearrayed into the sublibrary
BES 1824 containing 4100 putative expression clones. Note:
Due to a cloning artefact caused by the kit, in most cases
the SalI site is NOT present, as well as the SalI Adapter
used for cloning. To excise the insert, restriction sites
upstream SalI should be used (e.g. BamHI). Average insert
size is 1 kb. Library generation and sequencing was
granted in context of GABI; data are also accessible at
https://gabi.rzpd.de"

Query Match 87.0%; Score 17.4; DB 7; Length 893;
Best Local Similarity 89.5%; Pred. No. 8.8e+02;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGGUC 19
Db 564 AAATCTTCATAGGGTCG 546
|||||

RESULT 4
CNS06GLD
LOCUS CNS06GLD 975 bp DNA linear GSS 30-NOV-2001
DEFINITION T3 end of clone AS0AA005E07 of library AS0AA from strain CLIB 533
of Saccharomyces bayanus, genomic survey sequence.
ACCESSION AL397847
VERSION AL397847.1 GI:12150829
KEYWORDS GSS.

```

Saccharomyces bayanus
Saccharomyces bayanus
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
1 (bases 1 to 1003)
Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Boivin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Illorent,B.,
Malpartuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekla,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
PDBS Lett. 487 (1), 3-12 (2000)
20584711
PUBMED
11152876
2 (bases 1 to 1003)
Bon,E., Neuveglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,
Aigle,M. and Durrens,P.
Genomic exploration of the hemiascomycetous yeasts: 5.
Saccharomyces bayanus var. uvarum
PDBS Lett. 487 (1), 37-41 (2000)
20584715
PUBMED
11152880
3 (bases 1 to 1003)
Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
Location/Qualifiers
1..1003
/organism="Saccharomyces bayanus"
/mol_type="genomic DNA"
/strain="CLIB 533"
/variety="uvarum"
/db_xref="taxon:4931"
/clone="AS0AA026F08"
/clone_lib="AS0AA"
/notes="end : T3"
complement(<3..>988)
/notes="similar to Saccharomyces cerevisiae ORF YHR172w [
SPC97 : spindle pole body component]
1 putative frameshift(s)"
/evidence=not_experimental

misc_feature
Query Match 87.0%; Score 17.4; DB 9; Length 1003;
Best Local Similarity 89.8%; Pred. No. 8.9e+02;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAATTCTTTTAAAGGCGUC 19
|||||
DB 909 AAATTCTTGTAAGGTCG 927
|||||

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 1003;
Best Local Similarity 89.8%; Pred. No. 8.9e+02;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAATTCTTTTAAAGGCGUC 19
|||||
DB 909 AAATTCTTGTAAGGTCG 927
|||||

RESULT 6
CNS06J81/c
LOCUS
DEFINITION
T7 end of clone AS0AA026F08 of library AS0AA from strain CLIB 533
of Saccharomyces bayanus, genomic survey sequence.
ACCESSION
VERSION
AL401255.1 GI:12158665

```

KEYWORDS      GSS.
SOURCE        Saccharomyces bayanus
ORGANISM      Saccharomyces bayanus
REFERENCE     Eukaryote; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS       Saccharomycetales; Saccharomycetaceae; Saccharomyces.
1 (bases 1 to 1028)
REFERENCE     Souciet,J.L., Aigle,M., Bon,E., Artiguenave,F., Blandin,G.,
AUTHORS       Bolotin-Fukuhara,M., Mon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
TITLE        Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL      FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE      20584711
PUBMED       11152876
REFERENCE     2 (bases 1 to 1028)
AUTHORS       Bon,E., Neuveglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,
Aigle,M. and Durrens,P.
TITLE        Genomic exploration of the hemiascomycetous yeasts: 5.
Saccharomycetes bayanus var. uvarum
JOURNAL      FEBS Lett. 487 (1), 37-41 (2000)
MEDLINE      20584715
PUBMED       11152880
REFERENCE     3 (bases 1 to 1028)
AUTHORS       Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
COMMENT       2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
segrif@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES     Location/Qualifiers
             1..1028
             /organism="Saccharomyces bayanus"
             /mol_type="genomic DNA"
             /strain="CLIB 533"
             /variety="uvarum"
             /db_xref="taxon:4931"
             /clone_lib="AS0AA026F08"
             /clone_ref="AS0AA"
             /notes="end : T7"
             misc_feature
             <8..>1027
             /notes="similar to Saccharomyces cerevisiae ORF YHR172w [
             SP297 ; spindle pole body component ]
             2 putative frameshift(s)"
             /evidence=not_experimental
ORIGIN
Query Match      87.0%; Score 17.4; DB 9; Length 1028;
Best Local Similarity 89.5%; Pred. No. 8.9e+02;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAATCTTTTATAAGGUCG 19
|||||
Db 820 AAATCTTGATAGGTCG 802
|||||
RESULT 7
AG265169/c      85 bp      DNA      linear      GSS 22-JUL-2003
LOCUS          Lotus corniculatus var. japonicus DNA, clone:Lj762e07_sfi, genomic
DEFINITION     survey sequence.
ACCESSION      AG265169

```

```

VERSION        AG265169.1      GI:26665008
KEYWORDS       GSS.
SOURCE         Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM       Lotus corniculatus var. japonicus
REFERENCE     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS       Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.
1
REFERENCE     Sato,S., Nakamura,Y. and Tabata,S.
AUTHORS       Lotus japonicus TAC End sequences
TITLE         Published Only in Database (2002)
JOURNAL       2 (bases 1 to 85)
REFERENCE     Sato,S.
AUTHORS       Direct Submission
TITLE         Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,
JOURNAL       The First Laboratory for Plant Gene Research; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:ssato@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/,
Tel:81-438-52-3935(ex.2336), Fax:81-438-52-3934)
FEATURES     Location/Qualifiers
             1..85
             /organism="Lotus corniculatus var. japonicus"
             /mol_type="genomic DNA"
             /strain="Miyakoima MG-20"
             /variety="japonicus"
             /db_xref="taxon:34305"
             /clone_lib="genomic TAC library"
             /clone_ref="Lj762e07_sfi"
             /note="VECTOR: DTLTAC7-synonym: Lotus japonicus"
ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 85;
Best Local Similarity 85.0%; Pred. No. 1.4e+03;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAATCTTTTATAAGGUCGA 20
|||||
Db 36 AAATCTTAATATGGTCGA 17
|||||
RESULT 8
BB009906/c      293 bp      mRNA      linear      EST 22-JUN-2000
LOCUS          RIKEN full-length enriched, 10 day neonate skin Mus
DEFINITION     musculus cDNA clone 4732491N16.3', mRNA sequence.
ACCESSION      BB009906
VERSION        BB009906.1      GI:8130263
KEYWORDS       Mus musculus (house mouse)
SOURCE         Mus musculus
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 293)
REFERENCE     Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
AUTHORS       Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
Takahashi,P., Tomimaga,N., Toya,T., Tsunoda,Y., Watahiki,A.,
Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,
Yokota,T., Yoshida,K., Yoshiaki,A., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)

```

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S.,
 Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.:
 Thermosensitization and thermoinactivation of thermolabile enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
 Okazaki, Y., and Hayashizaki, Y.:
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.:
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.

FEATURES

Location/Qualifiers
 1..293
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4732491N16"
 /sex="mixed"
 /tissue_type="skin"
 /dev_stage="10 days neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 10 day neonate
 skin"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGATCTCGAGTCTTAATTAATCCCCCCCC
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. cDNA went through one round of normalization
 to Rot = 10.0 and subtraction to Rot = 100.0. Second
 strand cDNA was prepared with the primer adapter of
 sequence [5' GAGAGAGATCTCGAGTCTTAATTAATCCCCCCCC
 3']. cDNA was cloned into the XhoI and BamHI sites.
 Vector: a modified pBluescript KS(+) after bulk excision
 from Lambda FLC I"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 293;
 Best Local Similarity 85.0%; Pred. No. 1.6e+03;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGUGCA 20

Db 59 AAATCTTTATAGGUGCA 40

RESULT 9

LOCUS

AV258784 RIKEN full-length enriched, adult male testis (DH10B) Mus
 DEFINITION
 musculus cDNA clone 430401P07 3' similar to M19654 Mouse
 testis-specific phosphoglycerate kinase (pgk-2) mRNA, mRNA
 sequence.

ACCESSION

AV258784

VERSION

AV258784.1 GI:6246243

KEYWORDS

EST,

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 293)
 Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,
 Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,
 Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I.,
 Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
 Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K.,
 Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y.,
 Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomimaga, N.,
 Teunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A.,
 Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Konno, H., et al. 1999)
 Unpublished (1999)

TITLE

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
 Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and
 Hayashizaki, Y.:
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
 Okazaki, Y. and Hayashizaki, Y.:
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.:
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.

FEATURES

source

Location/Qualifiers
 1..299
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4930401P07"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, adult male testis
 (DH10B)"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGATCTCGAGTCTTAATTAATCCCCCCCC 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5'
 GAGAGAGATCTCGAGTCTTAATTAATCCCCCCCC 3'], cDNA
 was cloned into the XhoI and BamHI sites. Vector: a
 modified pBluescript KS(+) after bulk excision from Lambda
 FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI."

ORIGIN

Query Match 84.0%; Score 16.8; DB 1; Length 299;
 Best Local Similarity 85.0%; Pred. No. 1.6e+03;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGUGCA 20

```

Db      6  AAATCTTTATAGGGGTAGA 25
|||||
|||||

RESULT 10
CA856765      407 bp      mRNA      linear      EST 17-DEC-2002
LOCUS      PESToac41h07.v1 Plasmodium falciparum 3D7 gametocyte cDNA library
DEFINITION      Plasmodium falciparum 3D7 cDNA 5', mRNA sequence.
ACCESSION      CA856765
VERSION      CA856765.1  GI:27159521
KEYWORDS      EST.
SOURCE      Plasmodium falciparum 3D7
ORGANISM      Plasmodium falciparum 3D7
REFERENCE      1  (bases 1 to 407)
AUTHORS      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
TITLE      Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D.,
JOURNAL      Marra,M., Hillier,L., Martin,J., Wyllie,T., Dante,M., Theising,B.,
COMMENT      Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I.,
Teagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A.,
Maguire,L., Richey,J., Watkins,J., Kennedy,S., Levinso,D.,
Waterston,R., Wilson,R. and Sibley,D.
WashU Plasmodium EST Project
Contact: L. David Sibley
WashU Plasmodium EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library was constructed by R. Haywood. DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: L. David Sibley
(sibley@bcm.wustl.edu), Washington University
Seq primer: -40UP from Gibco.
Location/Qualifiers
FEATURES
source
1. .407
/organism="Plasmodium falciparum 3D7"
/mol_type="mRNA"
/db_xref="taxon:36329"
/lab_host="DH10B (GenesHog, Invitrogen, Inc.)"
/clone_lib="Plasmodium falciparum 3D7 gametocyte cDNA
library"
/notes="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:
XhoI; The library was constructed by R Haywood. cDNAs were
synthesized from gametocyte poly(A)+ RNA by oligo d(T)
priming, size-selected and directionally cloned into the
EcoRI (5' end) and XhoI (3' end) sites of the Uni-ZAP XR
lambda vector (Stratagene). The primary library was mass
excised as phagemid using the ExAssist helper phage
(Stratagene). Clones were mass excised using the ExAssist
helper phage (Stratagene), the phagemids were precipitated
with PEG 8000 and electroporated into DH10B cells. Clone
Phagemid DNA was electroporated into DH10B cells. Clone
Availability: David Sibley, Washington University."
ORIGIN
Query Match      84.0%; Score 16.8; DB 6; Length 407;
Best Local Similarity 85.0%; Pred. No. 1.6e+03;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1  AAATCTTTATAGGGGUCCA 20
|||||
Db      162 AAATTTTAAAGGTCGA 181
|||||
|||||

RESULT 11
BI926889      462 bp      mRNA      linear      EST 18-OCT-2001
LOCUS      EST546778 tomato flower, buds 0-3 mm Lycopersicon esculentum cDNA
DEFINITION      clone cTOA31E1 5' end, mRNA sequence.
ACCESSION      BI926889
VERSION      BI926889.1  GI:16235836
KEYWORDS      EST.
SOURCE      Lycopersicon esculentum (tomato)
ORGANISM      Lycopersicon esculentum
REFERENCE      1  (bases 1 to 462)
AUTHORS      van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
JOURNAL      Uterback,T., Van Aken,S., Romning,C.M., Nierman,W., Fraser,C.M.,
COMMENT      Martin,G.B., Giovannoni,J.J. and Tankaleley,S.D.
Generation of ESTs from tomato flower tissue, 0-3 mm buds (2001)
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
FEATURES
source
1. .462
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOA31E1"
/tissue_type="flower"
/dev_stage="0-3mm buds"
/clone_lib="tomato flower, buds 0-3 mm"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."
ORIGIN
Query Match      84.0%; Score 16.8; DB 4; Length 462;
Best Local Similarity 85.0%; Pred. No. 1.6e+03;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1  AAATCTTTATAGGGGUCCA 20
|||||
Db      356 AAATGTTTATAGGTCCTA 375
|||||
|||||

RESULT 12
BP527047/c      494 bp      mRNA      linear      EST 28-SEP-2004
LOCUS      BP527047
DEFINITION      BP527047 MAT001 Nicotiana tabacum cDNA clone BY11812, mRNA
sequence.
ACCESSION      BP527047
VERSION      BP527047.1  GI:52830774
KEYWORDS      EST.
SOURCE      Nicotiana tabacum (common tobacco)
ORGANISM      Nicotiana tabacum
REFERENCE      1  (bases 1 to 494)
AUTHORS      Matsuo,K., Tashiro,G., Horiguchi,T., Demura,T. and Fukuda,H.
JOURNAL      Profiling growth-phase dependent gene expression of tobacco BY-2
COMMENT      cells by comprehensive microarray analysis
Unpublished (2003)
Contact: Ken Matsuo
Morphogenesis Research Group
RIKEN Plant Science Center
1-7-2 Suehirocho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9575
Fax: 81-45-503-9573

```

Email: by2@psc.riken.go.jp, URL: <http://mrq.psc.riken.go.jp/strc/>
 The cDNA library was constructed from mRNA isolated from lag (9 h),
 lag (72 h) and stationary (7 days) old BY-2 cells.
 Seq primer: M13 forward.

FEATURES

source
 Location/Qualifiers
 1..494
 /organism="Nicotiana tabacum"
 /mol_type="mRNA"
 /cultivar="Bright Yellow No.2"
 /db_xref="taxon:4097"
 /clone="BY11812"
 /cell_line="BY-2"
 /clone_lib="MAT001"
 /notes="Vector: pGEM-T easy; primer: M13 forward; mRNA
 obtained from lag, log and stationary phase cells"

ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 494;
 Best Local Similarity 85.0%; Pred. No. 1.6e+03;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGGUCCA 20
 |||||
 Db 492 AAATCTTTAGAGGTCTCA 473

RESULT 13

CE435953 508 bp DNA linear GSS 27-SEP-2003
 LOCUS tigr-gss-dog-17000335866781 Dog Library Canis familiaris genomic.
 DEFINITION genomic survey sequence.

ACCESSION CE435953

VERSION CE435953.1 GI:36713767

KEYWORDS GSS.

SOURCE Canis familiaris (dog)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 508)

AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.

TITLE The dog genome: survey sequencing and comparative analysis

JOURNAL Science 301 (5641), 1898-1903 (2003)

MEDLINE 22875432

PUBMED 14512627

COMMENT Contact: Kirkness EF

The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirknes@tigr.org
 Class: shotgun.

FEATURES

source
 Location/Qualifiers
 1..508
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_lib="Dog Library"
 /notes="Site 1: BstXI; Libraries were prepared from
 peripheral blood"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 508;
 Best Local Similarity 85.0%; Pred. No. 1.6e+03;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGGUCCA 20
 |||||
 Db 15 AAATCTTTATAGGTGGA 34

RESULT 14

LOCUS CR154100

DEFINITION

CR154100 572 bp DNA linear GSS 06-JUL-2004
 Forward strand read from insert in 3'HPRT insertion targeting and
 chromosome engineering clone MHPF347b24, genomic survey sequence.

ACCESSION CR154100

VERSION CR154100.1 GI:49932945

KEYWORDS GSS; genome survey sequence; MICR.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 572)

REFERENCE Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,
 Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,
 Rogers, J. and Bradley, A.

AUTHORS

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. <http://www.sanger.ac.uk/MICR>

FEATURES

Location/Qualifiers

1..572

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHPF347b24"

/clone_lib="MHPF"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 572;
 Best Local Similarity 85.0%; Pred. No. 1.7e+03;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAATCTTTATAGGGUCCA 20
 |||||

Db 209 AAATCTTTATAGGATCCA 228

RESULT 15

LOCUS CD865385

DEFINITION

CD865385 588 bp mRNA linear EST 11-JUL-2003
 AZO2.073N10R000926 AZO2 Triticum aestivum cDNA clone AZO2073N10,
 mRNA sequence.

ACCESSION CD865385

VERSION CD865385.1 GI:32549201

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 588)

AUTHORS Genoplante.

TITLE Genoplante, a major partnership french program in plant genomics

JOURNAL Unpublished (2003)

COMMENT Contact: Genoplante

Genoplante
 93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' (<http://www.genoplante.com>
 and <http://genoplante-info.infobiogen.fr>).

FEATURES

source
 Location/Qualifiers
 1..588
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="recital"
 /db_xref="taxon:4565"
 /clone="AZO2073N10"
 /tissue_type="root"
 /clone_lib="AZO2"

ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 588;
 Best Local Similarity 85.0%; Pred. No. 1.7e+03;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAATCTTTTATAGGUGCA 20
 ||| ||||| ||||| :||
 Db 157 AAAATCTTTATAGGCTA 176

Search completed: March 17, 2005, 11:07:55
 Job time : 1389.27 secs